

Copyright (c) 1993 - 2006 Bioceleration Ltd.

Result No.	Score	Query			DB ID	Description
		Match	Length	%		
1	56	94.9	13	3	AAV99544	Hepatitis
2	56	94.9	13	3	AAV99551	Hepatitis
3	51	86.4	13	3	AAV99543	Hepatitis
4	39	66.1	71	4	AAW83798	Human imm
5	39	66.1	212	8	ADK16827	Nanoarcha
6	39	66.1	256	2	AEY74133	Human pro
7	39	66.1	330	9	AEBA1438	L. pneumo
8	39	66.1	377	9	AEB38149	L. pneumo
9	39	66.1	485	6	ABU11642	Human MDD
10	39	66.1	650	4	ABY94795	Human pro
11	39	66.1	650	8	ABW82021	Tumour-as
12	38	64.4	697	8	ADSA44988	Bacterial
13	37	62.7	10	3	AAV99545	Hepatitis
14	37	62.7	10	3	AAV97499	Hepatitis
15	37	62.7	165	9	AEBA1594	L. pneumo
16	37	62.7	165	9	AEB38314	L. pneumo
17	37	62.7	249	8	ADX67657	Plant ful
18	37	62.7	766	2	AAAG66222	Sucrose-s
19	36	61.0	94	4	AAO10535	Human pol
20	36	61.0	98	8	ADG32168	Mutant B
21	36	61.0	347	7	ADP04283	Bacterial
22	36	61.0	357	8	ADX92641	Plant ful
23	36	61.0	459	7	ABM87610	Rice abio
24	36	61.0	617	3	AAV76048	Murine sk

XX The present sequence is one of a large number of peptidic inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The present peptide is active at nanomolar
 CC levels and could therefore be used to treat or prevent hepatitis C or a
 CC related condition

XX Revised record issued on 18-NOV-2004 : Missing qualifiers

XX Sequence 13 AA;

Query Match 94.9%; Score 56; DB 3; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.0023;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLIXCPXDLPYK 13
 | | | | | | | | | |
 Db 1 DELIXCPXDLPYK 13

RESULT 2

AA99551
 ID AAY99551 standard; peptide; 13 AA.

AC AAY99551;

DT 18-NOV-2004 (revised)

DT 11-SEP-2000 (first entry)

XX Hepatitis C virus NS3 protease peptidic inhibitor #95.

XX Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 5 /note= "beta-Cyclohexylalanine"

FT Modified-site 8 /note= "beta-Cyclohexylalanine"

FT Modified-site 13 /note= "beta-Cyclohexylalanine"

FT Modified-site 13 /note= "N-Acetyl-lysine"

XX WO200031129-A1.

XX 02-JUN-2000.

XX 24-NOV-1999; 99WO-EP009207.

XX 26-NOV-1998; 98GB-00025946.

XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX Pessi A, Ingallinella P, Bianchi E;

XX WPI; 2000-411933/35.
 DR Peptide compounds useful for treating or preventing hepatitis C
 XX infections bind to hepatitis C virus NS3 protease.
 XX Example 1; Page 24; 46pp; English.

XX The present sequence is one of a large number of peptidic inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease

XX Revised record issued on 18-NOV-2004 : Missing qualifiers

XX Sequence 13 AA;

Query Match 94.9%; Score 56; DB 3; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.0023;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLIXCPXDLPYK 13
 | | | | | | | | | |
 Db 1 DELIXCPXDLPYK 13

RESULT 3

AA99543
 ID AAY99543 standard; peptide; 13 AA.

AC AAY99543;

DT 18-NOV-2004 (revised)

DT 11-SEP-2000 (first entry)

XX Hepatitis C virus NS3 protease peptidic inhibitor #87.

XX Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 3 /note= "3,3-Diphenylalanine"

FT Modified-site 5 /note= "beta-Cyclohexylalanine"

FT Modified-site 8 /note= "beta-Cyclohexylalanine"

FT Modified-site 13 /note= "N-epsilon-acetyl-lysine"

XX WO200031129-A1.

XX 02-JUN-2000.

XX 24-NOV-1999; 99WO-EP009207.

PR 26-NOV-1998; 98GB-00025946.
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 PA Pessi A, Ingallinella P, Bianchi E;
 PI WPI; 2000-411933/35.
 DR Peptide compounds useful for treating or preventing hepatitis C
 PT infections bind to hepatitis C virus NS3 protease.
 XX Example 2; Page 32; 46pp; English.
 XX The present sequence is one of a large number of peptidic inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The present peptide is active at nanomolar
 CC levels and could therefore be used to treat or prevent hepatitis C or a
 CC related condition
 CC Revised record issued on 18-NOV-2004 : Missing qualifiers
 CC Sequence 13 AA;
 XX
 SQ

Query Match 86.4%; Score 51; DB 3; Length 13;
 Best Local Similarity 84.6%; Pred. No. 0.02;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DXLIXCPXDLPLYK 13
 | | | | | | | | | | | | |
 DB 1 DEXIXCPXDLPLYK 13

RESULT 4
 AAM83798
 ID AAM83798 standard; protein; 71 AA.
 XX
 AC AAM83798;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 XX Human immune/haematopoietic antigen SEQ ID NO:11391.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis.
 KW Homo sapiens.
 XX
 OS WO200157182-A2.
 XX
 PN 09-AUG-2001.
 PD
 XX
 XX 17-JAN-2001; 2001WO-US0001354.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180828P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 26-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 14-AUG-2000; 2000US-0226273P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 DR N-PSDB; AAK56579.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 XX Claim 11; SEQ ID NO 11391; 3071pp + Sequence Listing; English.
 PS
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX Sequence 71 AA;
 SQ

Query Match 66.1%; Score 39; DB 4; Length 71;
 Best Local Similarity 77.8%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPDLP 11
 |||||
 Db 11 LIFCPDLP 19

RESULT 5

ADK16827
 ID ADK16827 standard; protein; 212 AA.

XX
 AC ADK16827;

XX
 DT 06-MAY-2004 (first entry)

XX
 DE Nanoarchaeum equitans cancer-associated (CA) protein #389.

XX
 KW cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia.

XX
 OS Nanoarchaeum equitans.

XX
 PN WO2003093434-A2.

XX
 PD 13-NOV-2003.

XX
 PF 01-MAY-2003; 2003WO-US013699.

XX
 PR 01-MAY-2002; 2002US-0377447P.

XX
 PA (DIVE-) DIVERSA CORP.

XX
 PI Stetter KO, Waters E, Kretz K, Podar M, Richardson T;
 PI Noordewier M;

XX
 XX WPI; 2004-053041/05.

DR
 N-PSDB; ADK16826.

XX
 PT New recombinant cancer-associated genes, such as KCNU9, useful for
 PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
 PT cervical, or skin cancers, lymphomas, or leukemia.
 XX

PS Claim 64; SEQ ID NO 779; 251pp; English.

XX
 CC The invention comprises then amino acid and coding sequences of cancer-
 CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention
 CC also comprises the Nanoarchaeum equitans genome. The DNA and protein
 CC sequences of the invention are useful for diagnosing and treating cancer
 CC (e.g. carcinoma, lymphoma, or leukaemia). The present amino acid sequence
 CC represents a Nanoarchaeum equitans CA protein of the invention.
 XX

XX
 SQ Sequence 212 AA;

Query Match 66.1%; Score 39; DB 8; Length 212;
 Best Local Similarity 75.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY      6 CPXDLPLYK 13
DB      199 CPLDCPYK 206

RESULT 6
AAY74133
ID AAY74133 standard; protein; 256 AA.
XX
AC AAY74133;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor EST fragment derived protein #320.
XX
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX
OS Homo sapiens.
XX
PN DE19820190-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1998; 98DE-01020190.
XX
PR 28-APR-1998; 98DE-01020190.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-G21386/54.
XX
DR N-PSDB; AA252963.
XX
PT New human nucleic acid sequences from pancreatic tumors, and related
XX
PS proteins.
XX
PS Claim 23; Page 442; 502pp; German.
XX
CC This invention describes novel polypeptides and their encoding nucleic
XX
CC acids derived from human pancreatic tumor tissue which have cytostatic
XX
CC activity. The sequences are also useful in producing pharmaceutical
XX
CC compositions for treatment of pancreatic tumors. AAY73914-Y74252
XX
CC represent protein fragments encoded by the human pancreatic tumor cDNA
XX
CC library derived expressed sequence tag (EST) sequences represented in
XX
CC AA252858-Z53014
XX
SQ Sequence 256 AA;

Query Match 66.1%; Score 39; DB 2; Length 256;
Best Local Similarity 53.8%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 DXLIXCPDLPLYK 13
DB      162 DINVGCPIDIVYK 174

RESULT 7
AEB41438
ID AEB41438 standard; protein; 330 AA.
XX
AC AEB41438;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila protein SEQ ID NO 5770.
XX
KW detection; infection; Antibacterial; Vaccine.
XX
OS Legionella pneumophila.

QY      1 DXLIXCPDLPLYK 12
DB      131 DMLKSCPKSLPY 142

RESULT 8
AEB38149
ID AEB38149 standard; protein; 377 AA.
XX
AC AEB38149;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila protein SEQ ID NO 2481.
XX
KW detection; infection; Antibacterial; Vaccine.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.

```

```

XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP ) INST PASTEUR.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rueniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
DR WPI; 2005-388305/40.
XX
PT New genome of Legionella pneumophila Paris strain and derived
XX
PT polypeptides, useful for detection or identification of the strain and
XX
PT for treatment and prevention of infections.
XX
PS Claim 3; SEQ ID NO 5770; 660pp; English.
XX
CC The invention relates to an isolated or purified nucleotide sequences (I)
XX
CC from Legionella pneumophila Paris strain. (I), and their related
XX
CC sequences or fragments, are useful as primers and probes for detection
XX
CC and amplification, including differentiation between the Paris and
XX
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
XX
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
XX
CC specific antibodies (Ab), also used for detection/identification of
XX
CC Legionella, and some (I), specifically those involved in synthesis of
XX
CC surface proteins, are targets for identification of inhibitors. (II), or
XX
CC vectors that contain (I), are useful as vaccines and immunogenic
XX
CC compositions, for treatment and prevention of infections by L.
XX
CC pneumophila. The present sequence represents the amino acid sequence of a
XX
CC L. pneumophila protein.
XX
SQ Sequence 330 AA;

Query Match 66.1%; Score 39; DB 9; Length 330;
Best Local Similarity 58.3%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 DXLIXCPDLPLY 12
DB      131 DMLKSCPKSLPY 142

RESULT 8
AEB38149
ID AEB38149 standard; protein; 377 AA.
XX
AC AEB38149;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila protein SEQ ID NO 2481.
XX
KW detection; infection; Antibacterial; Vaccine.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.

```

PA (INSP) INST PASTEUR.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
 PI Rueniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
 PI Jarraud S;
 XX WPI; 2005-388305/40.
 DR New genome of Legionella pneumophila Paris strain and derived
 XX polypeptides, useful for detection or identification of the strain and
 PT for treatment and prevention of infections.
 PT
 XX Claim 3; SEQ ID NO 2481; 660pp; English.
 XX
 CC The invention relates to an isolated or purified nucleotide sequences (I)
 CC from Legionella pneumophila Paris strain, (II), and their related
 CC sequences or fragments, are useful as primers and probes for detection
 CC and amplification, including differentiation between the Paris and
 CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
 CC (hybrid) polypeptides (III). (II) are also useful for preparation of
 CC specific antibodies (Ab), also used for detection/identification of
 CC Legionella, and some (I), specifically those involved in synthesis of
 CC surface proteins, are targets for identification of inhibitors (II), or
 CC vectors that contain (I), are useful as vaccines and immunogenic
 CC compositions, for treatment and prevention of infections by L.
 CC pneumophila. The present sequence represents the amino acid sequence of a
 CC L. pneumophila protein.
 XX
 SQ Sequence 377 AA;
 Query Match 66.1%; Score 39; DB 9; Length 377;
 Best Local Similarity 58.3%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLPY 12
 DB 178 DMLKSCPKSLPY 189
 RESULT 9
 ABU11642
 ID ABU11642 standard; protein; 485 AA.
 XX
 AC ABU11642;
 XX
 DT 12-FEB-2003 (first entry)
 DE Human MDDT polypeptide SEQ ID 589.
 XX
 XX MDDT; human; disease detection and treatment molecule polypeptide;
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
 KW haemostatic; nephrotropic; anti-anaemic; antipruritic; hepatotropic;
 KW gene therapy; protein replacement therapy; cell proliferative disorder;
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
 KW psoriasis; hepatitis.
 XX
 OS Homo sapiens.
 XX
 XX WO200279449-A2.
 FN
 XX 10-OCT-2002.
 PD
 XX 27-MAR-2002; 2002WO-US009944.
 FF
 XX 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
 PI Daugherty SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstein EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX WPI; 2003-058431/05.
 DR N-PSDB; ABX34632.
 DR
 XX New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 PT or hepatitis.
 XX Claim 27; SEQ ID NO 589; 339pp + Sequence Listing; English.
 PS
 XX This invention describes a novel disease detection and treatment molecule
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
 CC anti-anaemic, antipruritic and hepatotropic activity. The polynucleotides
 CC and the polypeptides of the invention can be used for gene therapy,
 CC protein replacement therapy and are useful for treating a variety of
 CC diseases or conditions. These polypeptides or polynucleotides are
 CC particularly useful for diagnosing, treating or preventing cell
 CC proliferative disorders (e.g. cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma or sarcoma), anaemia, Crohn's
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded
 CC by ABU11450-ABU11845, described in the disclosure of the invention. NOTE:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 485 AA;
 Query Match 66.1%; Score 39; DB 6; Length 485;
 Best Local Similarity 53.8%; Pred. No. 1.4e+02;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLPYK 13
 DB 394 DINVGCPIDLVIK 406
 RESULT 10
 AAB94795
 ID AAB94795 standard; protein; 650 AA.
 XX
 AC AAB94795;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:15918.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 FN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-00116126.
 PF
 XX

PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 PT
 PT
 XX
 XX
 PS Claim 8; SEQ ID NO 15918; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AA03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AA03166 to AAH13628 and
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 XX
 SQ Sequence 650 AA;
 Query Match 66.1%; Score 39; DB 4; Length 650;
 Best Local Similarity 53.8%; Pred. No. 1.9e+02;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLVYK 13
 Db 391 DINVGCPIDL VYK 403
 RESULT 11
 ID ABM82021
 XX ABM82021 standard; protein; 650 AA.
 AC
 XX ABM82021;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Tumour-associated antigenic target (TAT) polypeptide PRO82912, SEQ:5211.
 DE
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.
 XX
 XX Homo sapiens.

XX WO2004030615-A2.
 XX
 XX 15-APR-2004.
 XX
 XX 29-SEP-2003; 2003WO-US028547.
 XX
 XX 02-OCT-2002; 2002US-0414971P.
 XX (GETH) GENENTECH INC.
 XX
 XX Wu TD, Zhang Z, Zhou Y;
 XX
 XX WPI; 2004-347921/32.
 DR N-PSDB; ACN40433.
 DR
 PT New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 PT
 XX
 PS Claim 12; SEQ ID NO 5211; 7273pp; English.
 XX
 CC The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention
 XX
 XX
 SQ Sequence 650 AA;
 Query Match 66.1%; Score 39; DB 8; Length 650;
 Best Local Similarity 53.8%; Pred. No. 1.9e+02;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLVYK 13
 Db 391 DINVGCPIDL VYK 403
 RESULT 12
 ADS44988
 ID ADS44988 standard; protein; 697 AA.
 XX
 XX ADS44988;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX Bacterial polypeptide #23418.
 DE
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX

OS Bacteria.
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 23418; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 697 AA;
 Query Match 64.4%; Score 38; DB 8; Length 697;
 Best Local Similarity 54.5%; Pred. No. 3.1e+02;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXDLP 11
 Db 29 DVLLTCPLDCP 39
 RESULT 13
 AAY99545
 ID AAY99545 standard; peptide; 10 AA.
 XX
 AC AAY99545;
 XX
 DT 18-NOV-2004 (revised)
 DT 11-SEP-2000 (first entry)
 XX
 XX Hepatitis C virus NS3 protease peptidic inhibitor #89.

KW Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Misc-difference 2 /note= "D-form residue"
 FT Modified-site 5
 FT Modified-site 8 /note= "beta-Cyclohexylalanine"
 FT Modified-site 8 /note= "beta-Cyclohexylalanine"
 XX
 FN WO200031129-A1.
 XX
 XX 02-JUN-2000.
 PD
 XX
 XX 24-NOV-1999; 99WO-EP009207.
 PF
 XX 26-NOV-1998; 98GB-00025946.
 PR
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 PA
 XX Pessi A, Ingallinella P, Bianchi E;
 PI WPI; 2000-411933/35.
 DR
 XX
 XX Peptide compounds useful for treating or preventing hepatitis C
 PT infections bind to hepatitis C virus NS3 protease.
 PT
 XX Example 2; Page 32; 46pp; English.
 PS
 XX The present sequence is one of a large number of peptidic inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The present peptide is active at nanomolar
 CC levels and could therefore be used to treat or prevent hepatitis C or a
 CC related condition
 CC
 CC Revised record issued on 18-NOV-2004 : Missing qualifiers
 CC
 XX Sequence 10 AA;
 SQ
 Query Match 62.7%; Score 37; DB 3; Length 10;
 Best Local Similarity 90.0%; Pred. No. 5.7;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXDL 10
 Db 1 DELIXCPXDL 10
 RESULT 14
 AAY97499
 ID AAY97499 standard; peptide; 10 AA.
 XX
 AC AAY97499;
 XX

DT 11-SEP-2000 (first entry)
 XX Hepatitis C virus NS3 protease peptidic inhibitor #64.
 DE Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT
 FT Misc-difference 2 /note= "D-form residue"
 FT
 FT Modified-site 5 /note= "beta-Cyclohexylalanine"
 FT
 FT Modified-site 8 /note= "beta-Cyclohexylalanine"
 FT
 FT
 XX WO200031129-A1.
 PN
 XX
 PD 02-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-EP009207.
 XX
 PR 26-NOV-1998; 98GB-00025946.
 XX
 XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
 PA
 XX Pessi A, Ingallinella P, Bianchi E;
 PI
 XX WPI; 2000-411933/35.
 DR
 XX
 XX Peptide compounds useful for treating or preventing hepatitis C
 PT infections bind to hepatitis C virus NS3 protease.
 PT
 XX Disclosure; Page 17; 46pp; English.
 PS
 XX The present sequence is one of the most preferred examples from a large
 CC number of peptides that were synthesised as potential inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The compounds can be used at micromolar and
 CC nanomolar levels to treat or prevent hepatitis C or a related condition
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 62.7%; Score 37; DB 3; Length 10;
 Best Local Similarity 90.0%; Pred. No. 5.7;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDL 10
 DB 1 DELIXCPXDL 10
 RESULT 15
 AEB41594
 ID AEB41594 standard; protein; 165 AA.
 XX

AC AEB41594;
 XX
 DT 08-SEP-2005 (first entry)
 XX
 DE L. pneumophila protein SEQ ID NO 5926.
 XX
 KW detection; infection; Antibacterial; Vaccine.
 XX
 OS Legionella pneumophila.
 XX
 PN WO2005049642-A2.
 XX
 PD 02-JUN-2005.
 XX
 PF 23-SEP-2004; 2004WO-IB003578.
 XX
 PR 21-NOV-2003; 2003FR-00013687.
 XX
 PA (INSP) INST PASTEUR.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
 PI Rueniock C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
 PI Jarraud S;
 XX
 DR WPI; 2005-388305/40.
 XX
 XX New genome of Legionella pneumophila Paris strain and derived
 PT polypeptides, useful for detection or identification of the strain and
 PT for treatment and prevention of infections.
 PT
 XX Claim 3; SEQ ID NO 5926; 660pp; English.
 PS
 XX The invention relates to an isolated or purified nucleotide sequences (I)
 CC from Legionella pneumophila Paris strain. (I), and their related
 CC sequences or fragments, are useful as primers and probes for detection
 CC and amplification, including differentiation between the Paris and
 CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
 CC (hybrid) polypeptides (II). (II) are also useful for preparation of
 CC specific antibodies (Ab), also used for detection/identification of
 CC Legionella, and some (I), specifically those involved in synthesis of
 CC surface proteins, are targets for identification of inhibitors. (II), or
 CC vectors that contain (I), are useful as vaccines and immunogenic
 CC compositions, for treatment and prevention of infections by L.
 CC pneumophila. The present sequence represents the amino acid sequence of a
 CC L. pneumophila protein.
 XX
 SQ Sequence 165 AA;
 Query Match 62.7%; Score 37; DB 9; Length 165;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 6 CPXDLPY 12
 DB 74 CPDDIPY 80
 Search completed: March 15, 2006, 19:44:23
 Job time : 343.935 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:44:52 ; Search time 26.5652 Seconds
(without alignments)
47.085 Million cell updates/sec

Title: US-09-856-886b-103
Perfect score: 59
Sequence: 1 DXLIXCPXDLPLYK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: Pirl: *
2: Pirl: *
3: Pirl: *
4: Pirl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	72.9	730	2 T33041	hypothetical prote
2	38	64.4	697	2 C75525	molybdopter in oxid
3	37	62.7	822	2 S71493	sucrose synthase (
4	26	61.0	89	2 S52551	thionin variant Th
5	36	61.0	133	2 S52554	thionin variant Th
6	36	61.0	182	2 D70859	hypothetical prote
7	36	61.0	294	2 T33588	hypothetical prote
8	36	61.0	303	2 H70389	formate dehydrogen
9	36	61.0	317	2 S35282	eaA protein - phag
10	36	61.0	605	2 T35047	hypothetical prote
11	36	61.0	795	2 T34468	hypothetical prote
12	36	61.0	1286	2 B71413	hypothetical prote
13	35	59.3	135	2 S52555	thionin variant Th
14	35	59.3	135	2 S52552	thionin variant Th
15	35	59.3	192	2 B81911	probable molybdopt
16	35	59.3	210	2 C64614	molybdopter in bios
17	35	59.3	224	2 H82395	transcriptional regu
18	35	59.3	235	2 D71901	hypothetical prote
19	35	59.3	304	2 B96494	protein F7F22.14 l
20	35	59.3	368	2 H81059	phosphoserine amin
21	35	59.3	368	2 F81816	phosphoserine tran
22	35	59.3	521	2 T49355	related to protein
23	35	59.3	896	2 S48761	trehalose-phosphat
24	34	57.6	132	2 G71555	hypothetical prote
25	34	57.6	154	2 T35544	probable quinolina
26	34	57.6	194	2 B86073	hypothetical prote
27	34	57.6	194	2 D91226	Molybdopter in-guan
28	34	57.6	194	2 S40803	molybdopter in-guan
29	34	57.6	234	2 T05999	hypothetical prote

30 34 57.6 250 2 S36769
31 34 57.6 401 2 E96570
32 34 57.6 604 2 T08211
33 34 57.6 934 2 B29838
34 34 57.6 1190 2 T38636
35 33.5 56.8 98 2 JA0059
36 33 55.9 120 2 T02586
37 33 55.9 276 1 JU0205
38 33 55.9 282 2 T17398
39 33 55.9 395 2 D70441
40 33 55.9 398 1 B64383
41 33 55.9 442 2 C81978
42 33 55.9 442 2 C81034
43 33 55.9 449 2 T49646
44 33 55.9 478 2 A45285
45 33 55.9 486 2 A41537

ALIGNMENTS

RESULT 1

T33041
hypothetical protein T23E1.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33041
R:Scheet, P.; Kemp, K.
submitted to the EMBL Data Library, February 1998
A:Description: The sequence of C. elegans cosmid T23E1.
A:Reference number: Z21268
A:Accession: T33041
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-730.<SCH>
A:Cross-references: UNIPROT:O45204; UNIPARC:UPI00000830F9; EMBL:AF047665; PIDN:AAC04457.1
A:Experimental source: strain Bristol N2; clone T23E1
C:Genetics:
A:Gene: CESP:T23E1.2
A:Map position: 4
A:Introns: 20/2; 87/3, 139/2, 186/2, 233/2, 494/2, 546/2, 561/3, 583/1, 639/1

Query Match 72.9%; Score 43; DB 2; Length 730;
Best Local Similarity 60.0%; Pred. No. 4.5;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IXCPXDLPLYK 13
: ||| |||
DB 687 VDCPDVPEYK 696

RESULT 2

C75525
molybdopter in oxidoreductase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004
C:Accession: C75525
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-697 <WHI>
A:Cross-references: UNIPROT:Q9RXB8; UNIPARC:UPI00000D3CB7; GB:AE001899; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0397
A:Map position: 1

```
Query Match      64.4%; Score 38; DB 2; Length 697;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPL 11
Db 29 DVLLTCPLDLP 39

RESULT 3
S71493
sucrose synthase (EC 2.4.1.13) - beet
C:Species: Beta vulgaris (beet)
C>Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 09-Jul-2004
C:Accession: S71493; S71494
R:Hesse, H.; Willmitzer, L.
Plant Mol. Biol. 30, 863-872, 1996
A>Title: Expression analysis of a sucrose synthase gene from sugar beet (Beta vulgaris L.)
A:Reference number: S71493; MUID:96270366; PMID:8633746
A:Accession: S71493
A:Molecule type: mRNA
A:Residues: 1-822 <HES>
A:Cross-references: UNIPROT:Q42652; UNIPARC:UPI00001753C5; EMBL:X81974
R:Hesse, H.; Willmitzer, L.
submitted to the EMBL Data Library, September 1994
A:Description: Molecular cloning of a mitochondrial isoform of Cysteine Synthase from Arabidopsis thaliana
A:Reference number: S71494
A:Accession: S71494
A:Molecule type: mRNA
A:Residues: 'AG', 59-822 <HEM>
A:Cross-references: UNIPARC:UPI000013627B; EMBL:X81974; NID:gl488569; PIDN:CAA57499.1; F
C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:280-755/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match      62.7%; Score 37; DB 2; Length 822;
Best Local Similarity 58.3%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPL 12
Db 214 DYLIKLPADTPV 225

RESULT 4
S52551
thionin variant Thi2Ca7 - Abyssinian crambe
C:Species: Crambe abyssinica (Abyssinian crambe)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 31-Dec-2004
C:Accession: S52551
R:Schrader-Fischer, G.; Apel, K.
Mol. Gen. Genet. 245, 380-389, 1994
A>Title: Organ-specific expression of highly divergent thionin variants that are distinct
A:Reference number: S52545; MUID:95115690; PMID:7816048
A:Accession: S52551
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-89 <SCH>
A:Cross-references: UNIPARC:UPI00001766ED
C:Superfamily: hordothionin precursor

Query Match      61.0%; Score 36; DB 2; Length 89;
Best Local Similarity 62.5%; Pred. NO. 11;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 20 CPANLPYE 27

RESULT 5
S52554
thionin variant Thi2Call - Abyssinian crambe
```

```
C:Species: Crambe abyssinica (Abyssinian crambe)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 31-Dec-2004
C:Accession: S52554
R:Schrader-Fischer, G.; Apel, K.
Mol. Gen. Genet. 245, 380-389, 1994
A>Title: Organ-specific expression of highly divergent thionin variants that are distinct
A:Reference number: S52545; MUID:95115690; PMID:7816048
A:Accession: S52554
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-133 <SCH>
A:Cross-references: UNIPROT:Q9S975; UNIPARC:UPI00001766EE

Query Match      61.0%; Score 36; DB 2; Length 133;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 63 CPNLPYE 70

RESULT 6
D70859
hypothetical protein Rv3033 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70859
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70859
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-182 <COL>
A:Cross-references: UNIPROT:O53280; UNIPARC:UPI00000D1187; GB:AL021287; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3033

Query Match      61.0%; Score 36; DB 2; Length 182;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 IXCPXDLPL 11
Db 61 VTCPSDLP 68

RESULT 7
T33588
hypothetical protein C40A11.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33588
R:Maggi, L.; Goela, D.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid C40A11.
A:Reference number: Z21374
A:Accession: T33588
A:Molecule type: DNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-294 <MAG>
A:Cross-references: UNIPROT:Q9TZB0; UNIPARC:UPI000007EEDC; EMBL:AF099914; PIDN:AAC68758.1
A:Experimental source: strain Bristol N2; clone C40A11
C:Genetics:
A:Gene: CESP:C40A11.4
A:Map position: 2
A:Introns: 33/2; 74/3; 103/3; 183/3; 222/1
```

Query Match 61.0%; Score 36; DB 2; Length 294;
 Best Local Similarity 57.1%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 1 DXLIXCPXD--LPY 12
 |||||
 DB 195 DELINCPEDFFFPY 208
 |||||

RESULT 8
 H70389
 formate dehydrogenase beta subunit - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Oct-2004
 C:Accession: H70389
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: H70389
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-303 <AQF>
 A:Cross-references: UNIPROT:O67147; UNIPARC:UPI00000564F6; GB:AE000720; NID:G2983529; PI
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: fdhH
 C:Superfamily: formate dehydrogenase, iron-sulfur subunit; ferredoxin 2[4Fe-4S] homology
 F:136-196/Domain: ferredoxin 2[4Fe-4S] homology <PER>

Query Match 61.0%; Score 36; DB 2; Length 303;
 Best Local Similarity 55.6%; Pred. No. 37;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LIXCPXDLP 11
 |||||
 DB 150 LVGCPFDIP 158
 |||||

RESULT 9
 S35282
 eA protein - phage P22
 C:Species: phage P22
 C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
 C:Accession: S35282
 R:Wulff, D.L.; Ho, Y.S.; Powers, S.; Rosenberg, M.
 Mol. Microbiol. 9, 261-271, 1993
 A:Title: The int genes of bacteriophages P22 and lambda are regulated by different mecha
 A:Reference number: S35280; MUID:94018622; PMID:8412679
 A:Accession: S35282
 A:Molecule type: DNA
 A:Residues: 1-317 <WUL>
 A:Cross-references: UNIPROT:Q03544; UNIPARC:UPI0000003A7E; EMBL:L06296; NID:G215287; PI
 C:Genetics:
 A:Gene: eaa
 C:Superfamily: phage P22 eA protein

Query Match 61.0%; Score 36; DB 2; Length 317;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 CPXDLPY 12
 |||||
 DB 185 CPAELPY 191
 |||||

RESULT 10
 T35047
 hypothetical protein SC4G2.12c - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C:Accession: T35047; T30206
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z21566
 A:Accession: T35047
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-605 <SEE>
 A:Cross-references: UNIPROT:O86684; UNIPARC:UPI00000DADA6; EMBL:AL031371; PIDN:CAA20549.1
 A:Experimental source: strain A3(2)
 R:Bedford, D.J.; Laity, C.; Butner, M.J.
 J. Bacteriol. 177, 4681-4689, 1995
 A:Title: Two genes involved in the phase-variable phi C31 resistance mechanism of Streptoc
 A:Reference number: Z20777; MUID:95370146; PMID:7642495
 A:Accession: T30206
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 304-605 <BED>
 A:Cross-references: UNIPARC:UPI00000BEAA5; EMBL:L37531; NID:G576537; PID:G576541; PIDN:A
 C:Genetics:
 A:Gene: SCORDB:SC4G2.12c

Query Match 61.0%; Score 36; DB 2; Length 605;
 Best Local Similarity 85.7%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PXDLPLYK 13
 |||||
 DB 365 PADLPLYK 371
 |||||

RESULT 11
 T34468
 hypothetical protein ZK770.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T34468
 R:Maggi, L.; Gattung, S.; Bartko, L.
 submitted to the EMBL Data Library, April 1997
 A:Description: The sequence of C. elegans cosmid ZK770.
 A:Reference number: Z21530
 A:Accession: T34468
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795 <MAG>
 A:Cross-references: UNIPROT:O01635; UNIPARC:UPI0000129183; EMBL:U97404; PIDN:AAB93309.1;
 A:Experimental source: strain Bristol N2; clone ZK770
 C:Genetics:
 A:Gene: CESP:ZK770.1
 A:Map position: 1
 A:Introns: 18/1; 84/2; 125/3; 198/3; 257/3; 337/3; 378/3; 424/3; 477/2; 510/1; 563/3; 634/

Query Match 61.0%; Score 36; DB 2; Length 795;
 Best Local Similarity 75.0%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 CPXDLPYK 13
 |||||
 DB 626 CPVDDPYK 633
 |||||

RESULT 12
 B71413
 hypothetical protein dl3525w - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Variety: columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
 C:Accession: B71413
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
 avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech

erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: B71413
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1286 <BEV>
A;Cross-references: UNIPROT:Q23351; UNIPARC:UPI00000A522B; GB:Z97337; NID:g2244829; PID:
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: Arabidopsis thaliana hypothetical protein dl325w

Query Match 61.0%; Score 36; DB 2; Length 1286;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPL 11
||:||||
Db 570 DALLGCPHLP 580

RESULT 13
S52555
thionin variant Thi2Ca12 - Abyssinian crambe
C;Species: Crambe abyssinica (Abyssinian crambe)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 31-Dec-2004
C;Accession: S52555
R;Schrader-Fischer, G.; Apel, K.
Mol. Gen. Genet. 245, 380-389, 1994
A;Title: Organ-specific expression of highly divergent thionin variants that are distinct
A;Reference number: S52545; MUID:95115690; PMID:7816048
A;Accession: S52555
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-135 <SCH>
A;Cross-references: UNIPROT:Q9S974; UNIPARC:UPI000009D9CC
C;Superfamily: hordothionin precursor

Query Match 59.3%; Score 35; DB 2; Length 135;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CPXDLPY 12
||:||||
Db 65 CPNLPY 71

RESULT 14
S52552
thionin variant Thi2Ca9 - Abyssinian crambe
C;Species: Crambe abyssinica (Abyssinian crambe)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 31-Dec-2004
C;Accession: S52552
R;Schrader-Fischer, G.; Apel, K.
Mol. Gen. Genet. 245, 380-389, 1994
A;Title: Organ-specific expression of highly divergent thionin variants that are distinct
A;Reference number: S52545; MUID:95115690; PMID:7816048
A;Accession: S52552
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-135 <SCH>
A;Cross-references: UNIPROT:Q9S977; UNIPARC:UPI00000A5CCD
C;Superfamily: hordothionin precursor

Query Match 59.3%; Score 35; DB 2; Length 135;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CPXDLPY 12
||:||||
Db 65 CPNLPY 71

RESULT 15
B81911
probable molybdopterin-guanine dinucleotide biosynthesis protein A NMA1417 [imported] -
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-Oct-2004
C;Accession: B81911
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel]
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Murgall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: B81911
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-192 <PAR>
A;Cross-references: UNIPROT:Q9JUA5; UNIPARC:UPI0000012F354; GB:AL162755; GB:AL157959; NID:
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1417
C;Superfamily: molybdopterin guanylyltransferase Moba

Query Match 59.3%; Score 35; DB 2; Length 192;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
||:||||
Db 93 DWLLVPCDWPY 104

Search completed: March 15, 2006, 19:50:12
Job time : 29.5652 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rnp**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:
<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher. When submitting sequence search results for scanning into IRW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:34:25 ; Search time 164.478 Seconds
(without alignments)
55.763 Million cell updates/sec

Title: US-09-856-886B-103
Perfect score: 59
Sequence: 1 DKLXCPXDLPLYK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	72.9	222	2	Q86FL9_CAEEL	Q86f19 caenorhabdi
2	43	72.9	666	2	Q4QA30_LEIMA	Q4qa30 leishmania
3	42	71.2	283	2	Q7PR13_ANOGA	Q7pr13 anopheles g
4	40	67.8	583	2	Q585Q4_9TRYP	Q585q4 trypanosoma
5	40	67.8	704	2	Q72WK8_DESVH	Q72wk8 desulfovibr
6	40	67.8	1145	2	Q7SIS6_9RETR	Q7sis6 simian foam
7	39	66.1	185	2	Q88LD6_PSEPK	Q88ld6 pseudomonas
8	39	66.1	190	2	Q5NGX2_FRATT	Q5ngx2 francisella
9	39	66.1	212	2	Q74N20_NAMEQ	Q74n20 nanoarchaeu
10	39	66.1	223	2	Q95323_HUMAN	Q95323 homo sapien
11	39	66.1	330	2	Q5X8B2_LEGPA	Q5x8b2 legionella
12	39	66.1	339	2	Q4FMF2_9RICK	Q4fmf2 candidatus
13	39	66.1	350	2	Q5WZR2_LEGPL	Q5wzr2 legionella
14	39	66.1	350	2	Q5ZYN4_LEIONA	Q5zyn4 leionella
15	39	66.1	408	2	Q9BSU4_HUMAN	Q9bsu4 homo sapien
16	39	66.1	423	1	ECTB_STRAW	Q82914 streptomyce
17	39	66.1	426	2	Q7TT12_MOUSE	Q7tt12 mus musculu
18	39	66.1	465	2	Q9NPR1_HUMAN	Q9npr1 homo sapien
19	39	66.1	613	2	Q96HM5_HUMAN	Q96hm5 homo sapien
20	39	66.1	637	2	Q91X11_MOUSE	Q91x11 mus musculu
21	39	66.1	639	2	Q6P2X3_XENTR	Q6p2x3 xenopus tro
22	39	66.1	640	2	Q7ZWS1_XENLA	Q7zws1 xenopus lae
23	39	66.1	650	2	Q9H877_HUMAN	Q9h877 homo sapien
24	39	66.1	650	2	Q96G46_HUMAN	Q96g46 homo sapien
25	39	66.1	2539	2	Q9J7C1_PRRSV	Q9j7c1 porcine rep
26	38	64.4	215	2	Q8F5X9_LEPIN	Q8f5x9 leptospira
27	38	64.4	215	2	Q7Q282_LEPIC	Q7q282 leptospira
28	38	64.4	345	2	Q4NDP8_9MICC	Q4ndp8 arthrobacte
29	38	64.4	697	2	Q9RXB8_DEIRA	Q9rxb8 deinococcus
30	38	64.4	1434	2	Q6PN63_BUNYW	Q6pn63 ngari virus
31	37	62.7	174	2	Q6NF07_CORDI	Q6nf07 corynebacte

32	37	62.7	192	2	Q72BN1_DESVH	Q72bni desulfovibr
33	37	62.7	305	2	Q8QND6_9PHYC	Q8qnd6 ectocarpus
34	37	62.7	321	2	Q552C9_DICDI	Q552c9 dictyosteli
35	37	62.7	356	2	Q9RDY6_LEGPN	Q9rdy6 legionella
36	37	62.7	543	2	Q9W645_CHICK	Q9w645 gallus gall
37	37	62.7	599	2	Q8DV61_STRMU	Q8dv61 streptococc
38	37	62.7	747	2	Q4WJ12_ASPFU	Q4wj12 aspergillus
39	37	62.7	766	1	SUSY_BETVU	Q42652 beta vulgar
40	37	62.7	822	2	Q94G60_BETVU	Q94g60 beta vulgar
41	37	62.7	2106	2	Q4ILV9_GIBZE	Q4ilv9 gibberella
42	36	61.0	51	2	Q9S978_CRAAB	Q9s978 crambe abys
43	36	61.0	98	2	Q62T02_BACLD	Q62t02 bacillus li
44	36	61.0	130	2	Q9S975_CRAAB	Q9s975 crambe abys
45	36	61.0	132	2	Q5DBU4_SCHJA	Q5dbj4 schistosoma

ALIGNMENTS

RESULT 1
Q86FL9_CAEEL
ID Q86FL9_CAEEL PRELIMINARY; PRT; 222 AA.
AC Q86FL9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein T23E1.3.
GN ORFNames=T23E1.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AF047665; AAP13767.1; -, Genomic_DNA.
DR Ensembl; T23E1.3; Caenorhabditis elegans.
DR WormBase; WBGene0020730; T23E1.3.
DR WormPep; T23E1.3; CE33838.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 222 AA; 26444 MW; 1588F999C61D061D CRC64;

Query Match 72.9%; Score 43; DB 2; Length 222;
Best Local Similarity 60.0%; Pred. No. 5.3;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IXCPXDLPYK 13
: ||| :|||
Db 179 VDCPQDVPYK 188

RESULT 2
Q4QA30_LEIMA
ID Q4QA30_LEIMA PRELIMINARY; PRT; 666 AA.
AC Q4QA30;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF25.0650;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S.; Murphy L.; Ivens A.C.; Berriman M.; Blackwell J.;
RA Smith D.; Collins M.; Foster N.; Harris D.; Oliver K.; O'Neil S.;

```

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005264; CAJ04848.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 666 AA; 68968 MW; 7F092DBB65CB4CC CRC64;
Query Match 72.9%; Score 43; DB 2; Length 666;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 DXLIXCPXDLPL 11
Db 621 DVLLCPADLP 631

RESULT 3
ID Q7PRI3 ANOGA PRELIMINARY; PRT; 283 AA.
AC Q7PRI3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000023293.
GN ORFNames=ENSANGG0000020716;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL "Anopheles gambiae re-annotation.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008849; EAA07155.3; -; Genomic DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR Pfam; PF01607; CBM14; 2.
DR SMART; SM00494; ChEBD2; 1.
DR PROSITE; PS00940; CHIT_BIND II; 2.
SQ SEQUENCE 283 AA; 30920 MW; 556B99F0BE777230 CRC64;
Query Match 71.2%; Score 42; DB 2; Length 283;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 DXLIXCPXDLPL 12
Db 86 DLTIACPTDPY 97

RESULT 4
ID Q585Q4 9TRYP PRELIMINARY; PRT; 583 AA.
AC Q585Q4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Tb927.6.1560;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

```

```

OX NCBI_TaxID=5691;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shallem J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC013353; AAX79710.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 583 AA; 65925 MW; A26D29B890E3FF7F CRC64;
Query Match 67.8%; Score 40; DB 2; Length 583;
Best Local Similarity 54.5%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 DXLIXCPXDLPL 11
Db 221 DLLLSCPTDFP 231

RESULT 5
ID Q72WK8 DESVH PRELIMINARY; PRT; 704 AA.
AC Q72WK8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycosyl transferase, group 1/2 family protein.
GN OrderedLocuNames=DVUA0081;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OG Plasmid pDV.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
[1]
RN NUCLEOTIDE SEQUENCE.
RP PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolony J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidesen T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uitterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017286; AAS94377.1; -; Genomic_DNA.
DR TIGR; DVUA0081; -.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR01296; Glyco_trans_1.
DR InterPro; IPR01173; Glyco_trans_2.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR Complete proteome.
SQ SEQUENCE 704 AA; 77934 MW; 9272EF83D31E8AA CRC64;
Query Match 67.8%; Score 40; DB 2; Length 704;
Best Local Similarity 41.7%; Pred. No. 68;

```

```

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DXLIXCPDLPY 12
Db 372 EAAVACPEDVPY 383

RESULT 6
Q7S1S6_9RETR PRELIMINARY; PRT; 1145 AA.
AC Q7S1S6;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polymerase.
DE Name=pol;
GN Simian foamy virus-orangutan.
OS Simian foamy viruses; Retroviridae; Spumavirus.
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.
OX NCBI_TaxID=221703;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SFVora; TISSUE=Blood;
RX MEDLINE=22742031; PubMed=12857929;
RY DOI=10.1128/JVI.77.15.8584-8587.2003;
RA Verschoor E.J., Langenhuijzen S., van den Engel S., Niphuis H.,
RA Warren K.S., Heeney J.L.;
RT "Structural and evolutionary analysis of an orangutan foamy virus.";
RL J. Virol. 77:8584-8587(2003).
DR EMBL; AJ544579; CAD67562.1; -; Genomic_DNA.
DR HSP; P03355; I163.
DR MEROPS; A09_001; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:006310; P:DNA recombination; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001641; Peptidase_A9_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF03539; Spuma_A9Pase; 1.
DR PRINTS; PR00920; SPUMVIRPTASE.
DR PROSITE; PS00879; RNaseH_H; 1.
SQ SEQUENCE 1145 AA; 128989 MW; F7DF9F1B6C614469 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 1145;
Best Local Similarity 41.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPDLPY 12
Db 81 DVIILCPADVPW 92

RESULT 7
Q88LD6_PSEPK PRELIMINARY; PRT; 185 AA.
AC Q88LD6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Bacteriocin production protein.
GN OrderedLocNames=PPI999;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;

```

```

RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodeon R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.F.,
RA Hance I., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Moazzaz A., Utterback T.R., Holtzapfle E.K., Scanlan D., Tran K.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmli B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016781; AAN67613.1; -; Genomic_DNA.
DR TIGR; PPI999; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009403; P:toxin biosynthesis; IEA.
DR InterPro; IPR003825; Colicin_V.
DR Pfam; PF02674; Colicin_V; 1.
KW Complete proteome.
SQ SEQUENCE 185 AA; 19937 MW; 7C693923D42AEC7F CRC64;

Query Match 66.1%; Score 39; DB 2; Length 185;
Best Local Similarity 63.6%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LIXCPXDLPYK 13
Db 166 LISTPADLPFK 176

RESULT 8
Q5NGX2_FRATT PRELIMINARY; PRT; 190 AA.
AC Q5NGX2;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=FTT0704;
OS Francisella tularensis (subsp. tularensis).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;
OC Francisellaceae; Francisella.
OX NCBI_TaxID=119856;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCHU S4 / Schu 4;
RX PubMed=15640799; DOI=10.1038/ng1499;
RA Larsson P., Oyston P.C.F., Chain P., Chu M.C., Duffield M.,
RA Fuxelius H.-H., Garcia E., Haelettorp G., Johansson D., Isherwood K.E.,
RA Karp P.D., Larsson E., Liu Y., Michell S., Prior J., Prior R.,
RA Malfatti S., Sjoestedt A., Svensson K., Thompson N., Verges L.,
RA Wagg J.K., Wren B.W., Lindler L.E., Andersson S.G.E., Forsman M.,
RA Titball R.W.;
RT "The complete genome sequence of Francisella tularensis, the causative
RT agent of tularemia.";
RL Nat. Genet. 37:153-159(2005).
DR EMBL; AJ749949; CAG45337.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 190 AA; 20975 MW; EC81551495CCDC39 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 190;
Best Local Similarity 53.8%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLPYK 13
Db 26 DSTSICPTDLKYK 38

```

```

RESULT 9
Q74N20_NANEQ
ID Q74N20_NANEQ PRELIMINARY; PRT; 212 AA.
AC Q74N20;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE NEQ398.
GN OrderedLocusNames=NEQ398;
OC Archaea; Nanoarchaeum.
OX NCST_TaxID=160232;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PUBMed=14566062; DOI=10.1073/pnas.1735403100;
RA Waters E., Hohn M.J., Ahel I., Graham D.E., Adams M.D., Barnstead M.,
RA Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X.,
RA Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
RA Soell D., Stetter K.O., Short J.M., Noordwier M.;
RT "The genome of Nanoarchaeum equitans: insights into early archaeal
RT evolution and derived parasitism."
RL Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988(2003).
DR EMBL; AE017199; AAR39244.1; -; Genomic DNA.
DR GO; GO:006284; P:base-excision repair; IEA.
DR InterPro; IPR003265; Endo_3c.
DR Pfam; PF00730; HHH-GPD; 1.
DR SMART; SM00478; ENDO3C; 1.
DR KW Complete proteome.
SQ SEQUENCE 212 AA; 25093 MW; 1C04D18C007084F4 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 212;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
| | | | |
Db 199 CPLDCPYK 206

RESULT 10
O95323_HUMAN
ID O95323_HUMAN PRELIMINARY; PRT; 223 AA.
AC O95323;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barrow I.K.-P., Boguski M.S., Touchman J., Spencer F.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the synthesis of dihydrouridine, a modified
CC base found in the D-loop of most tRNAs (By similarity).
CC -!- COPACTOR: FAD (By similarity).
DR EMBL; AF031072; AAC72941.1; -; mRNA.
DR GO; GO:0050660; F:ADP binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008033; P:tRNA processing; IEA.
DR InterPro; IPR001269; Du synth.
DR Pfam; PF01207; Duf; 1.
DR PROSITE; PS01136; UPF0034; UNKNOWN 1.
DR KW FAD; Hypothetical protein; tRNA processing.
FT NON_TER 1
FT TER 223
SQ SEQUENCE 223 AA; 25573 MW; 9222D88E77BFE68 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 330;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
| | | | |
Db 131 DMLKSCPKSLPY 142

RESULT 11
O5X8B2_LEGPA
ID O5X8B2_LEGPA PRELIMINARY; PRT; 330 AA.
AC O5X8B2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=lpp0341;
OS Legionella pneumophila (strain Paris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=297246;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15467720; DOI=10.1038/ng1447;
RA Caralec C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
RA Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
RA Etienne J., Glaser P., Buchrieser C.;
RT "Evidence in the Legionella pneumophila genome for exploitation of
RT host cell functions and high genome plasticity."
RL Nat. Genet. 36:1165-1173(2004).
DR EMBL; CR628336; CAH1489.1; -; Genomic DNA.
DR Legioli; lpp0341; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR002523; Mg2+_transpCorA.
DR Pfam; PF01544; CorA; 1.
DR KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 330 AA; 38380 MW; FDA2ED20E303A778 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 330;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
| | | | |
Db 131 DMLKSCPKSLPY 142

RESULT 12
O4FMF2_9RICK
ID O4FMF2_9RICK PRELIMINARY; PRT; 339 AA.
AC O4FMF2;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Putative rhamnosyltransferase.
GN Name=wbba; ORFNames=SAR11_0820;
OS Candidatus Pelagibacter ubique HTCC1062.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC SAR11 cluster; Candidatus Pelagibacter.
OX NCBI_TaxID=335992;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HTCC1062;
RA Giovannoni S.J., Tripp H.J., Givan S.A., Podar M., Vergin K.L.,
RA Baptista D., Bibbs L., Eads J., Richardson T.H., Noordweir M.,
RA Rappe M.S., Short J., Carrington J.C., Mathur E.J.;
RT "Genome Streamlining in a Cosmopolitan Oceanic Bacterium."
RT Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000084; AA221637.1; -; Genomic DNA.
DR KW Transferase.

```

SQ SEQUENCE 339 AA; 40444 MW; 9FB7C2FD5CEB90DB CRC64;
 Query Match 66.1%; Score 39; DB 2; Length 339;
 Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 LIXCPXDLFPY 12
 DB 229 IIMCPVDIPY 238
 RESULT 13
 Q5WZR2 LEGPL PRELIMINARY; PRT; 350 AA.
 AC Q5WZR2; DOI=10.1038/ng1447;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=lp10319;
 OS Legionella pneumophila (strain Lens).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Legionellaceae; Legionella.
 OX NCBI_TaxID=297245;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15467720; DOI=10.1038/ng1447;
 RA Cazale C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
 RA Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
 RA Etienne J., Glaser P., Buchrieser C.;
 RT "Evidence in the Legionella pneumophila genome for exploitation of
 host cell functions and high genome plasticity."
 RL Nat. Genet. 36:1165-1173(2004).
 DR EMBL; CR628337; CAH14550.1; -; Genomic_DNA.
 DR LegioList; lp10319; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0046873; F:metal ion transporter activity; IEA.
 DR InterPro; IPR002523; Mg2+_transpCorA.
 DR Pfam; PF01544; CorA; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 350 AA; 40600 MW; 99E6AA0AEC6B9127 CRC64;
 Query Match 66.1%; Score 39; DB 2; Length 350;
 Best Local Similarity 58.3%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLFPY 12
 DB 151 DLLKSCPKSLPY 162
 RESULT 14
 Q5ZYU4 LEGPH PRELIMINARY; PRT; 350 AA.
 AC Q5ZYU4;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Magnesium and cobalt transport protein CorA.
 GN Name=corA; OrderedLocusNames=lp90267;
 OS Legionella pneumophila subsp. pneumophila (strain Philadelphia 1 /
 ATCC 33152).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Legionellaceae; Legionella.
 OX NCBI_TaxID=272624;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15448271; DOI=10.1126/science.1099776;
 RA Chien M., Morozova I., Shi S., Sheng H., Chen J., Gomez S.M.,
 RA Asammi G., Hill K., Nuara J., Feder M., Rinner J., Greenberg J.J.,
 RA Steshenko V., Park S.H., Zhao S.H., Replitskaya E., Edwards J.R.,
 RA Pampou S., Georgioud A., Chou I.-C., Iannuccilli W., Uiz M.E.,

Kim D.H., Geringer-Sameth A., Goldsberry C., Morozov P., Fischer S.G.,
 Segal G., Qu X., Rzhetsky A., Zhang P., Cayanis E., De Jong P.J.,
 Ju J., Kalachikov S., Shuman H.A., Russo J.J.;
 RT "The genomic sequence of the accidental pathogen Legionella
 pneumophila";
 RL Science 305:1966-1968(2004).
 DR EMBL; AB017354; AAU26374.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0046873; F:metal ion transporter activity; IEA.
 DR GO; GO:0030001; P:metal ion transport; IEA.
 DR InterPro; IPR002523; Mg2+_transpCorA.
 DR Pfam; PF01544; CorA; 1.
 KW Complete proteome.
 SQ SEQUENCE 350 AA; 40739 MW; 41A4823FBBDF85F7 CRC64;
 Query Match 66.1%; Score 39; DB 2; Length 350;
 Best Local Similarity 58.3%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDLFPY 12
 DB 151 DLLKSCPKSLPY 162
 RESULT 15
 Q9BSU4 HUMAN PRELIMINARY; PRT; 408 AA.
 AC Q9BSU4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE DUS3L protein.
 GN Name=DUS3L;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX TISSUE=Placenta;
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Madsen K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Smailus D.E.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos U.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Placenta;
 RG NIH MGC Project;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004549; AAH04549.1; -; mRNA.
 DR GO; GO:0050660; F:FAD binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008033; P:tRNA processing; IEA.
 DR InterPro; IPR001269; Du_synth.

```

DR InterPro; IPR003009; FMN_enzyme.
DR Pfam; PF01207; Dus; 1.
DR PROSITE; PS01136; UFF0034; UNKNOWN 1.
SQ SEQUENCE 408 AA; 46106 MW; 96931E7C0C377C89 CRC64;

Query Match      66.1%; Score 39; DB 2; Length 408;
Best Local Similarity 53.8%; Pred. NO. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLPYK 13
   | : ||| |||
Db 149 DINVGCPIDLVIK 161

```

Search completed: March 15, 2006, 19:49:21
Job time : 168.478 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:49:41 ; Search time 37.3043 Seconds
(without alignments)
28.811 Million cell updates/sec

Title: US-09-856-886B-103
Perfect score: 59
Sequence: 1 DKLXCPDLPYK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/1aa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pcp:*
5: /cgn2_6/ptodata/1/1aa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/1aa/baCkfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	62.7	766	1	US-08-553-436A-8
2	36	61.0	347	2	US-09-543-681A-4568
3	36	61.0	617	2	US-09-188-930-303
4	36	61.0	617	2	US-09-312-283C-303
5	35	59.3	325	2	US-09-702-705-1816
6	35	59.3	325	2	US-09-736-457-1816
7	35	59.3	325	2	US-09-671-325-1816
8	35	59.3	325	2	US-10-017-754-1816
9	35	59.3	401	2	US-09-631-863A-2
10	35	59.3	609	2	US-08-969-683A-67
11	35	59.3	609	2	US-09-308-207-67
12	35	59.3	619	2	US-09-489-039A-10209
13	35	59.3	896	2	US-09-538-092-142
14	34	57.6	159	2	US-09-621-976-4449
15	34	57.6	250	1	US-08-861-269-5
16	34	57.6	250	1	US-09-134-596-5
17	34	57.6	250	2	US-09-293-273-5
18	34	57.6	255	2	US-09-902-540-14178
19	34	57.6	490	2	US-09-328-352-4282
20	33	55.9	13	2	US-09-719-261B-9
21	33	55.9	64	2	US-09-540-236-2010
22	33	55.9	140	2	US-09-902-540-15651
23	33	55.9	162	2	US-09-270-767-46244
24	33	55.9	193	2	US-09-540-236-2160
25	33	55.9	344	2	US-09-902-540-12625
26	33	55.9	352	2	US-09-902-540-12606
27	33	55.9	371	2	US-09-489-039A-8780

28	55.9	384	2	US-09-270-767-43890	Sequence 43890, A
29	55.9	414	2	US-09-248-796A-20956	Sequence 20956, A
30	55.9	486	2	US-09-538-092-1269	Sequence 1269, Ap
31	55.9	486	2	US-09-949-016-6151	Sequence 6151, Ap
32	55.9	510	2	US-09-949-016-10733	Sequence 10733, A
33	55.9	677	2	US-09-345-473E-38	Sequence 38, Appl
34	55.9	677	2	US-09-862-027-38	Sequence 38, Appl
35	55.9	744	2	US-09-252-991A-19290	Sequence 19290, A
36	55.9	796	2	US-09-949-002-386	Sequence 386, App
37	55.9	802	2	US-09-949-002-512	Sequence 512, App
38	54.2	11	2	US-08-676-242-10	Sequence 10, Appl
39	54.2	16	2	US-09-461-061-11	Sequence 11, Appl
40	54.2	27	2	US-09-461-061-8	Sequence 8, Appl
41	54.2	50	2	US-09-621-976-5997	Sequence 5997, Ap
42	54.2	117	1	US-08-193-114B-1	Sequence 1, Appl
43	54.2	117	4	PCT-US92-06809-1	Sequence 1, Appl
44	54.2	123	2	US-09-461-061-18	Sequence 18, Appl
45	54.2	144	2	US-09-270-767-33267	Sequence 33267, A

ALIGNMENTS

RESULT 1
US-08-553-436A-8
; Sequence 8, Application US/08553436A
; Patent No. 5866790
; GENERAL INFORMATION:
; APPLICANT: HESSE, Holger
; APPLICANT: MULLER-ROBER, Bernd
; TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
; TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE
; TITLE OF INVENTION: CONCENTRATION
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553.436A
; FILING DATE: 17-NOV-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/01671
; FILING DATE: 20-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 4317596.1
; FILING DATE: 24-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-553-436A-8

Query Match 62.7%; Score 37; DB 1; Length 766;

Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 158 DYLIKLPADTPY 169

RESULT 2

US-09-543-681A-4568
; Sequence 4568, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4568
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4568

Query Match 61.0%; Score 36; DB 2; Length 347;
Best Local Similarity 70.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LIXCPXDLPY 12
Db 293 LINIPRDLPY 302

RESULT 3

US-09-188-930-303
; Sequence 303, Application US/09188930A
; Patent No. 6150502

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-303

Query Match 61.0%; Score 36; DB 2; Length 617;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 IXCPXDLPY 12
Db 405 VLCPDDAPY 413

RESULT 4

US-09-312-283C-303
; Sequence 303, Application US/09312283C
; Patent No. 6573095

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312.283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-303

Query Match 61.0%; Score 36; DB 2; Length 617;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 IXCPXDLPY 12
Db 405 VLCPDDAPY 413

RESULT 5

US-09-702-705-1816
; Sequence 1816, Application US/09702705
; Patent No. 6504010

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702.705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1816
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-1816

Query Match 59.3%; Score 35; DB 2; Length 325;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 78 CPDLPEK 85

RESULT 6

US-09-736-457-1816
; Sequence 1816, Application US/09736457
; Patent No. 6509448

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1816
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1816

Query Match 59.3%; Score 35; DB 2; Length 325;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 78 CPEDLPEK 85

RESULT 7
US-09-671-325-1816
; Sequence 1816, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1816
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-1816

Query Match 59.3%; Score 35; DB 2; Length 325;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 78 CPEDLPEK 85

RESULT 8
US-10-017-754-1816
; Sequence 1816, Application US/10017754
; Patent No. 6858204
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1816
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-754-1816

Query Match 59.3%; Score 35; DB 2; Length 325;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 78 CPEDLPEK 85

RESULT 9
US-09-631-863A-2
; Sequence 2, Application US/09631863A
; Patent No. 6809179
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Tumor-associated Antigen R11
; FILE REFERENCE: 12211aa
; CURRENT APPLICATION NUMBER: US/09/631,863A
; CURRENT FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 2.1
; SEQ ID NO 2
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-631-863A-2

Query Match 59.3%; Score 35; DB 2; Length 401;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 154 CPEDLPEK 161

RESULT 10
US-08-969-683A-67
; Sequence 67, Application US/08969683A
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; STREET: 1870 South Winton road
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A.

```

; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,683A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20873
; FILING DATE: 13-NOV-1997
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX:
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6136576e
; US-08-969-683A-67

Query Match 59.3%; Score 35; DB 2; Length 609;
Best Local Similarity 41.7%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 120 EALLSCSADTPY 131

RESULT 11
US-09-308-207-67
; Sequence 67, Application US/09308207
; Patent No. 6953684
; GENERAL INFORMATION:
; APPLICANT: MARIA DIAZ-TORRES ET AL.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; CITY: 1870 South Winton road
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/308,207
; FILING DATE: 13-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888

```

```

; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6953684e
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:
US-09-308-207-67

Query Match 59.3%; Score 35; DB 2; Length 609;
Best Local Similarity 41.7%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 120 EALLSCSADTPY 131

RESULT 12
US-09-489-039A-10209
; Sequence 10209, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10209
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-10209

Query Match 59.3%; Score 35; DB 2; Length 619;
Best Local Similarity 41.7%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 129 EALLSCSADTPY 140

RESULT 13
US-09-538-092-142
; Sequence 142, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqformatter Version 0.9
; SEQ ID NO 142
; LENGTH: 896
; TYPE: PRT

```

```

; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR074W
US-09-538-092-142

Query Match          59.3%; Score 35; DB 2; Length 896;
Best Local Similarity 54.5%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      3 LIXCPXDLPYK 13
      :| | | |
Db      16 IINCVTQLPYK 26

RESULT 14
US-09-621-976-4449
; Sequence 4449, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4449
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4449

Query Match          57.6%; Score 34; DB 2; Length 159;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXLIXCPXD 9
      :| | | |
Db      88 DSLIQCPID 96

RESULT 15
US-08-861-269-5
; Sequence 5, Application US/08861269
; Patent No. 5817494
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,269
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: 397581
; CLONE: GenBank
US-08-861-269-5

Query Match          57.6%; Score 34; DB 1; Length 250;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      3 LIXCPXDLPYK 13
      :| | | |
Db      40 IITGPADTPYK 50

Search completed: March 15, 2006, 19:51:22
Job time : 37.3043 secs

```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:50:26 ; Search time 128.304 Seconds
(without alignments)
42.335 Million cell updates/sec

Title: US-09-856-886b-103
Perfect score: 59
Sequence: 1 DXLIXCPXDLVPK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	74.6	1618	4	US-10-424-599-206388
2	44	74.6	1874	4	US-10-424-599-248103
3	40	67.8	65	4	US-10-424-599-280379
4	39	66.1	94	4	US-10-424-599-182908
5	39	66.1	279	4	US-10-306-762-143
6	39	66.1	279	4	US-10-306-762-144
7	39	66.1	279	4	US-10-306-762-145
8	39	66.1	279	4	US-10-306-762-146
9	39	66.1	279	4	US-10-306-762-147
10	39	66.1	279	4	US-10-306-762-154
11	39	66.1	423	4	US-10-156-761-1324
12	38	64.4	127	4	US-10-425-115-276428
13	38	64.4	697	4	US-10-369-493-23418
14	37	62.7	32	4	US-10-437-963-164031
15	37	62.7	208	4	US-10-425-114-38500
16	37	62.7	249	4	US-10-437-963-120142
17	37	62.7	319	4	US-10-437-963-1259832
18	36	61.0	72	4	US-10-425-115-259832
19	36	61.0	83	4	US-10-425-115-268667
20	36	61.0	146	4	US-10-425-115-349436
21	36	61.0	357	4	US-10-425-114-55305
22	36	61.0	360	4	US-10-424-599-227759
23	36	61.0	459	4	US-10-437-963-203846
24	36	61.0	617	3	US-09-866-050A-303
25	36	61.0	1286	3	US-09-998-027-3
26	36	61.0	1286	4	US-10-165-099-3
27	35.5	60.2	397	4	US-10-425-115-214400

28	35	59.3	36	4	US-10-425-115-269638	Sequence 269638,
29	35	59.3	50	3	US-09-764-847-752	Sequence 752, App
30	35	59.3	50	4	US-10-092-154-752	Sequence 752, App
31	35	59.3	73	4	US-10-424-599-253505	Sequence 253505,
32	35	59.3	103	4	US-10-425-114-42684	Sequence 42684, A
33	35	59.3	134	4	US-10-425-115-212967	Sequence 212967,
34	35	59.3	137	4	US-10-437-963-188525	Sequence 188525,
35	35	59.3	148	4	US-10-437-963-188527	Sequence 188527,
36	35	59.3	173	4	US-10-437-963-151558	Sequence 151558,
37	35	59.3	217	4	US-10-425-114-48346	Sequence 48346, A
38	35	59.3	224	3	US-09-882-227-100	Sequence 100, App
39	35	59.3	235	4	US-10-335-977-5500	Sequence 5500, Ap
40	35	59.3	236	4	US-10-335-977-5501	Sequence 5501, Ap
41	35	59.3	243	5	US-10-450-763-38435	Sequence 38435, A
42	35	59.3	243	6	US-11-097-143-7821	Sequence 7821, Ap
43	35	59.3	264	4	US-10-424-599-267358	Sequence 267358,
44	35	59.3	281	4	US-10-306-762-164	Sequence 164, App
45	35	59.3	296	4	US-10-186-886-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-424-599-206388
; Sequence 206388, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206388
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28395C.1.pep
US-10-424-599-206388

Query Match 74.6%; Score 44; DB 4; Length 1618;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLPY 12
DB 1438 DILVSCPSMPY 1449

RESULT 2
US-10-424-599-248103
; Sequence 248103, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 248103
; LENGTH: 1874
; TYPE: PRT

ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_66067C.1.pep
US-10-424-599-248103

Query Match 74.6%; Score 44; DB 4; Length 1874;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 819 DILVSCPSPMEY 830

RESULT 3
US-10-424-599-280379
Sequence 280379, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 280379
LENGTH: 65
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_95204C.1.pep
US-10-424-599-280379

Query Match 67.8%; Score 40; DB 4; Length 65;
Best Local Similarity 60.0%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 IXCPXDLPYK 13
Db 53 IDCPSDYPYR 62

RESULT 4
US-10-424-599-182908
Sequence 182908, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 182908
LENGTH: 94
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_13617C.1.pep
US-10-424-599-182908

Query Match 66.1%; Score 39; DB 4; Length 94;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 7 CPKDLPHK 14

RESULT 5
US-10-306-762-143
Sequence 143, Application US/10306762
Publication No. US20030187220A1
GENERAL INFORMATION:
APPLICANT: Park, Frances
APPLICANT: Gajiwala, Ketan S.
APPLICANT: Buchanan, Sean Grant
APPLICANT: Sauder, J. Michael
TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
TITLE OF INVENTION: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
FILE REFERENCE: 52498-20011.00
CURRENT APPLICATION NUMBER: US/10/306,762
CURRENT FILING DATE: 2003-04-16
PRIOR APPLICATION NUMBER: US 60/334,132
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 242
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 143
LENGTH: 279
TYPE: PRT
ORGANISM: H. sapiens (13528705)
US-10-306-762-143

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPYK 13
Db 91 DINVGCPIDLVPYK 103

RESULT 6
US-10-306-762-144
Sequence 144, Application US/10306762
Publication No. US20030187220A1
GENERAL INFORMATION:
APPLICANT: Park, Frances
APPLICANT: Gajiwala, Ketan S.
APPLICANT: Buchanan, Sean Grant
APPLICANT: Sauder, J. Michael
TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
TITLE OF INVENTION: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
FILE REFERENCE: 52498-20011.00
CURRENT APPLICATION NUMBER: US/10/306,762
CURRENT FILING DATE: 2003-04-16
PRIOR APPLICATION NUMBER: US 60/334,132
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 242
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 144
LENGTH: 279
TYPE: PRT
ORGANISM: H. sapiens (14249957)
US-10-306-762-144

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPYK 13
Db 91 DINVGCPIDLVPYK 103

RESULT 7
US-10-306-762-145

```

; Sequence 145, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; PRIOR FILING DATE: 2003-04-16
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 279
; TYPE: PRT
; ORGANISM: H. sapiens (24308187)
US-10-306-762-145

```

```

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 DXLIXCPXDLVPYK 13
DB 91 DINVGCPIDLVYK 103

```

```

RESULT 8
US-10-306-762-146
; Sequence 146, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; PRIOR FILING DATE: 2003-04-16
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146
; LENGTH: 279
; TYPE: PRT
; ORGANISM: H. sapiens (10436076)
US-10-306-762-146

```

```

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 DXLIXCPXDLVPYK 13
DB 91 DINVGCPIDLVYK 103

```

```

RESULT 9
US-10-306-762-147
; Sequence 147, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael

```

```

; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 279
; TYPE: PRT
; ORGANISM: H. sapiens (22052882)
US-10-306-762-147

```

```

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 DXLIXCPXDLVPYK 13
DB 91 DINVGCPIDLVYK 103

```

```

RESULT 10
US-10-306-762-154
; Sequence 154, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154
; LENGTH: 279
; TYPE: PRT
; ORGANISM: M. musculus (21450077)
US-10-306-762-154

```

```

Query Match 66.1%; Score 39; DB 4; Length 279;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 DXLIXCPXDLVPYK 13
DB 91 DINVGCPIDLVYK 103

```

```

RESULT 11
US-10-156-761-13924
; Sequence 13924, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29

```

; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13924
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13924

Query Match 66.1%; Score 39; DB 4; Length 423;
Best Local Similarity 72.7%; Pred. No. 90;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LIXCPXDLPLYK 13
||| |||||
Db 98 LILRPDLPLYK 108

RESULT 12

US-10-425-115-276428
; Sequence 276428, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 276428
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(127)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_183685C.1.pep
US-10-425-115-276428

Query Match 64.4%; Score 38; DB 4; Length 127;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPLY 12
||| |||||
Db 94 DGLLACAEELPY 105

RESULT 13

US-10-369-493-23418
; Sequence 23418, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-21(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23418
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Deinococcus radiodurans
US-10-369-493-23418

Query Match 64.4%; Score 38; DB 4; Length 697;
Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPLY 11
||| |||||
Db 29 DVLLTCPLDCP 39

RESULT 14

US-10-437-963-164031
; Sequence 164031, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164031
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62971C.1.pep
US-10-437-963-164031

Query Match 62.7%; Score 37; DB 4; Length 32;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPLY 12
||| |||||
Db 20 DFLIGCPSNLLY 31

RESULT 15

US-10-425-115-366223
; Sequence 366223, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366223
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9715C.1.pep

US-10-425-115-366223

Query Match 62.7%; Score 37; DB 4; Length 208;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels

```
Qy      1 DXLIXCPXDLP 11
        | : | | : |
Db     51 DCMLLCPVDMP 61
```

Search completed: March 15, 2006, 19:55:15
Job time : 129.304 secs

This Page Blank (uspto)

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_NA_Main) and **.rapbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:51:36 ; Search time 16.9565 Seconds
(without alignments)
21.944 Million cell updates/sec

Title: US-09-856-886B-103
Perfect score: 59
Sequence: 1 DLXCPXDLFPYK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues
Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/prodata/1/pubpaa/US08 NEW PUB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US06 NEW PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US07 NEW PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/PCT_NEW PUB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US09 NEW PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US10 NEW PUB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US11 NEW PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US60_NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	39	66.1	423	7	US-11-087-099-9700
2	38	64.4	311	7	US-11-096-568A-3250
3	38	64.4	311	7	US-11-096-568A-3251
4	38	64.4	317	7	US-11-096-568A-3249
5	38	64.4	318	7	US-11-096-568A-3248
6	37	62.7	156	7	US-11-096-568A-25076
7	37	62.7	158	7	US-11-098-686-11309
8	37	62.7	208	7	US-11-096-568A-15370
9	37	62.7	306	7	US-11-096-568A-15369
10	37	62.7	320	7	US-11-096-568A-15368
11	36	61.0	98	6	US-10-510-386-138
12	36	61.0	372	7	US-11-098-686-10144
13	35	59.3	204	6	US-10-467-657-4830
14	35	59.3	399	6	US-10-821-234-1163
15	35	59.3	423	7	US-11-087-099-527
16	35	59.3	426	6	US-10-467-657-4808
17	34	57.6	602	7	US-11-131-263-10
18	34	57.6	602	7	US-11-131-263-19
19	34	57.6	602	7	US-11-131-263-29
20	34	57.6	604	7	US-11-131-263-40
21	34	57.6	759	7	US-11-096-568A-29706
22	34	57.6	764	7	US-11-096-568A-29705
23	34	57.6	767	7	US-11-096-568A-29704
24	33	55.9	118	7	US-11-096-568A-25077
25	33	55.9	170	6	US-10-467-657-7228

26	33	55.9	230	7	US-11-096-568A-8424	Sequence 8424, Ap
27	33	55.9	271	7	US-11-096-568A-8423	Sequence 8423, Ap
28	33	55.9	337	7	US-11-096-568A-5739	Sequence 5739, Ap
29	33	55.9	339	6	US-10-467-657-4318	Sequence 4318, Ap
30	33	55.9	408	6	US-10-763-712A-67	Sequence 67, Appl
31	33	55.9	442	6	US-10-467-657-1614	Sequence 1614, Ap
32	33	55.9	456	6	US-10-763-712A-66	Sequence 66, Appl
33	33	55.9	462	6	US-10-763-712A-60	Sequence 60, Appl
34	33	55.9	462	6	US-10-763-712A-68	Sequence 68, Appl
35	33	55.9	502	6	US-10-995-561-954	Sequence 954, Ap
36	33	55.9	795	6	US-11-242-553-14	Sequence 14, Appl
37	32	54.2	57	7	US-11-096-568A-27126	Sequence 27126, A
38	32	54.2	143	7	US-10-914-848-3	Sequence 3, Appl
39	32	54.2	165	6	US-11-096-568A-20990	Sequence 20990, A
40	32	54.2	282	7	US-11-096-568A-20989	Sequence 20989, A
41	32	54.2	341	7	US-11-098-686-11069	Sequence 11069, A
42	32	54.2	352	7	US-11-087-099-10661	Sequence 10661, A
43	32	54.2	475	7	US-11-054-515-952	Sequence 952, App
44	31	52.5	251	7	US-11-054-515-982	Sequence 982, App
45	31	52.5	251	7	US-11-054-515-982	Sequence 982, App

ALIGNMENTS

RESULT 1
US-11-087-099-9700
; Sequence 9700, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9700
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis MA-4680
US-11-087-099-9700

Query Match 56.1%; Score 39; DB 7; Length 423;
Best Local Similarity 72.7%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LIXCPXDLFPYK 13
|||
DB 98 LILRPDLFPYK 108

RESULT 2
US-11-096-568A-3250
; Sequence 3250, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3250
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(311)
; OTHER INFORMATION: Ceres Seq. ID no. 15171166
US-11-096-568A-3250

```
Query Match      64.4%; Score 38; DB 7; Length 311;
Best Local Similarity 45.5%; Pred. No. 5.9;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DXLIXCPXDLP 11
Db      153 DCMLLCPADMP 163

RESULT 3
US-11-096-568A-3251
; Sequence 3251, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3251
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(311)
; OTHER INFORMATION: Ceres Seq. ID no. 16625542
US-11-096-568A-3251

Query Match      64.4%; Score 38; DB 7; Length 311;
Best Local Similarity 45.5%; Pred. No. 5.9;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DXLIXCPXDLP 11
Db      153 DCMLLCPADMP 163

RESULT 4
US-11-096-568A-3249
; Sequence 3249, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3249
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 15171165
US-11-096-568A-3249

Query Match      64.4%; Score 38; DB 7; Length 317;
Best Local Similarity 45.5%; Pred. No. 6;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DXLIXCPXDLP 11
Db      159 DCMLLCPADMP 169

RESULT 5
US-11-096-568A-3248
```

```
; Sequence 3248, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3248
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(318)
; OTHER INFORMATION: Ceres Seq. ID no. 15171164
US-11-096-568A-3248

Query Match      64.4%; Score 38; DB 7; Length 318;
Best Local Similarity 45.5%; Pred. No. 6.1;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DXLIXCPXDLP 11
Db      160 DCMLLCPADMP 170

RESULT 6
US-11-096-568A-25076
; Sequence 25076, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 25076
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(156)
; OTHER INFORMATION: Ceres Seq. ID no. 12556308
US-11-096-568A-25076

Query Match      62.7%; Score 37; DB 7; Length 156;
Best Local Similarity 45.5%; Pred. No. 4.7;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DXLIXCPXDLP 11
Db      37 DCMLLCPDMP 47

RESULT 7
US-11-098-686-11309
; Sequence 11309, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
```

; PRIOR APPLICATION NUMBER: US 60/416,395
 ; PRIOR FILING DATE: 2002-10-04
 ; NUMBER OF SEQ ID NOS: 11433
 ; SOFTWARE: FASTSEQ For Windows Version 4.0
 ; SEQ ID NO 11309
 ; LENGTH: 158
 ; TYPE: PRT
 ; ORGANISM: Lawsonia intracellularis
 US-11-098-686-11309

Query Match 62.7%; Score 37; DB 7; Length 158;
 Best Local Similarity 63.6%; Pred. No. 4.7;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LIXCPXDLPLK 13
 DB 96 LIACPHDLWK 106

RESULT 8

US-11-096-568A-15370
 ; Sequence 15370, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 15370
 ; LENGTH: 208
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(208)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12344353
 US-11-096-568A-15370

Query Match 62.7%; Score 37; DB 7; Length 208;
 Best Local Similarity 45.5%; Pred. No. 6.2;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLP 11
 DB 51 DCMLLCFVDMP 61

RESULT 9

US-11-096-568A-15369
 ; Sequence 15369, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 15369
 ; LENGTH: 306
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(306)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12344352
 US-11-096-568A-15369

Query Match 62.7%; Score 37; DB 7; Length 306;

Best Local Similarity 45.5%; Pred. No. 9;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLP 11
 DB 149 DCMLLCFVDMP 159

RESULT 10

US-11-096-568A-15368
 ; Sequence 15368, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 15368
 ; LENGTH: 320
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(320)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12344351
 US-11-096-568A-15368

Query Match 62.7%; Score 37; DB 7; Length 320;
 Best Local Similarity 45.5%; Pred. No. 9.4;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXDLP 11
 DB 163 DCMLLCFVDMP 173

RESULT 11

US-10-510-386-138
 ; Sequence 138, Application US/10510386
 ; Publication No. US20050244922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Andersen, Jens Tonne
 ; APPLICANT: Clausen, Ib Groth
 ; APPLICANT: Jorgensen, Steen Troels
 ; APPLICANT: Olsen, Peter Bjarke
 ; APPLICANT: Rasmussen, Michael Dolberg
 ; TITLE OF INVENTION: Improved Bacillus Host Cell
 ; FILE REFERENCE: 10294-204-US
 ; CURRENT APPLICATION NUMBER: US/10/510,386
 ; CURRENT FILING DATE: 2004-10-04
 ; NUMBER OF SEQ ID NOS: 248
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 138
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bacillus licheniformis
 US-10-510-386-138

Query Match 61.0%; Score 36; DB 6; Length 98;
 Best Local Similarity 75.0%; Pred. No. 4.6;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 IXCPXDLP 11
 DB 41 INCPSDLP 48

RESULT 12

US-11-098-686-10144
 ; Sequence 10144, Application US/11098686

; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10144
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10144

Query Match 61.0%; Score 36; DB 7; Length 372;
Best Local Similarity 55.6%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 IXCPXDLPY 12
Db 39 VACPLDRPY 47

RESULT 13
US-10-467-657-4830
; Sequence 4830, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4830
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4830

Query Match 59.3%; Score 35; DB 6; Length 204;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDLPY 12
Db 105 DWLLIVPCDMFY 116

RESULT 14
US-10-821-234-1163
; Sequence 1163, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1163
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1163

Query Match 59.3%; Score 35; DB 6; Length 399;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 CPXDLPYK 13
Db 152 CPEDLPEK 159

RESULT 15
US-11-087-099-527
; Sequence 527, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 527
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor A3(2)
US-11-087-099-527

Query Match 59.3%; Score 35; DB 7; Length 423;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PXDLPYK 13
Db 102 PRDLPYK 108

Search completed: March 15, 2006, 19:55:51
Job time : 17.9565 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:34:00 ; Search time 259.565 Seconds
(without alignments)
16.927 Million cell updates/sec

Title: US-09-856-886B-104
Perfect score: 40
Sequence: 1 DLIIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 21:.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	92.5	10	AA999545	Aay99545 Hepatitis
2	37	92.5	10	AA997499	Aay97499 Hepatitis
3	37	92.5	13	AA999544	Aay99544 Hepatitis
4	37	92.5	13	AA999551	Aay99551 Hepatitis
5	36	90.0	1288	Adc42337	Adc42337 A. thalia
6	34	85.0	280	AA993918	AA993918 Human pro
7	34	85.0	332	AA993879	AA993879 Human pro
8	34	85.0	357	AA997274	AA997274 Novel cen
9	34	85.0	357	AD154589	AD154589 Novel hum
10	34	85.0	399	AD121120	AD121120 Novel hum
11	34	85.0	417	AA993029	AA993029 Human sec
12	34	85.0	437	AA994044	AA994044 Human pro
13	34	85.0	437	AA994058	AA994058 Human pro
14	34	85.0	437	AD121597	AD121597 Novel hum
15	34	85.0	440	AA992976	AA992976 Novel hum
16	34	85.0	440	AB910286	AB910286 Human cDN
17	34	85.0	440	AA997584	AA997584 Novel cen
18	34	85.0	440	AB966873	AB966873 Human pol
19	34	85.0	440	AD154899	AD154899 Novel hum
20	34	85.0	442	AA997291	AA997291 Novel cen
21	34	85.0	442	AD154606	AD154606 Novel hum
22	34	85.0	449	AB910269	AB910269 Human cDN
23	34	85.0	449	AB966856	AB966856 Human pol
24	33	82.5	816	AB992809	AB992809 Herbicide

25	33	82.5	914	4	ABG14451	Abg14451 Novel hum
26	33	82.5	1102	4	ABG07864	Abg07864 Novel hum
27	33	82.5	1102	4	ABG14252	Abg14252 Novel hum
28	32	80.0	9	3	AA999546	AA999546 Hepatitis
29	32	80.0	10	3	AA997514	AA997514 Hepatitis
30	32	80.0	10	3	AA997511	AA997511 Hepatitis
31	32	80.0	10	3	AA997512	AA997512 Hepatitis
32	32	80.0	10	3	AA997513	AA997513 Hepatitis
33	32	80.0	13	3	AA999543	AA999543 Hepatitis
34	32	80.0	31	2	AA999544	AA999544 Human bcl
35	32	80.0	71	4	AA983798	AA983798 Human imm
36	32	80.0	314	8	ADL05013	ADL05013 M. catarr
37	32	80.0	333	6	ABU35150	ABU35150 Protein e
38	32	80.0	697	8	ADS44988	ADS44988 Bacterial
39	31	77.5	10	3	AA999542	AA999542 Hepatitis
40	31	77.5	120	4	AA999550	AA999550 Hepatitis
41	31	77.5	120	4	AAO00146	AAO00146 Human pol
42	31	77.5	182	7	ADM25847	ADM25847 Hyperther
43	31	77.5	216	4	ABG10921	ABG10921 Novel hum
44	31	77.5	245	3	AA908334	AA908334 Arabidops
45	31	77.5	245	3	AA948220	AA948220 Arabidops

ALIGNMENTS

RESULT 1
AA999545
ID AA999545 standard; peptide; 10 AA.
XX
AC AA999545;
XX
DT 18-NOV-2004 (revised)
DT 11-SEP-2000 (first entry)
XX
DE Hepatitis C virus NS3 protease peptidic inhibitor #89.
XX
KW Hepatitis C virus; HCV; serine protease; virucide; infection;
KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
KW hepatocellular carcinoma.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Misc-difference 2 /note= "D-form residue"
FT Modified-site 5 /note= "beta-Cyclohexylalanine"
FT Modified-site 8 /note= "beta-Cyclohexylalanine"
WO200031129-A1.
02-JUN-2000.
XX
XX 24-NOV-1999; 99WO-EP009207.
XX
XX 26-NOV-1998; 98GB-00025946.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
PA Pessi A, Ingallinella P, Bianchi E;
PI WPI; 2000-411933/35.
XX
XX Peptide compounds useful for treating or preventing hepatitis C
XX infections bind to hepatitis C virus NS3 protease.
XX
XX Example 2; Page 32; 46pp; English.
XX
XX The present sequence is one of a large number of peptidic inhibitors of

CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The present peptide is active at nanomolar
 CC levels and could therefore be used to treat or prevent hepatitis C or a
 CC related condition
 CC
 CC Revised record issued on 18-NOV-2004 : Missing qualifiers
 CC
 CC Sequence 10 AA;

Query Match 92.5%; Score 37; DB 3; Length 10;
 Best Local Similarity 90.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
 | | | | |
 Db 1 DELIXCPXDL 10

RESULT 2
 AAY97499
 ID AAY97499 standard; peptide; 10 AA.
 AC AAY97499;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Hepatitis C virus NS3 protease peptidic inhibitor #64.
 XX
 KW Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.
 XX
 OS Synthetic.

Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Misc-difference 2 /note= "D-form residue"
 FT Modified-site 5 /note= "beta-Cyclohexylalanine"
 FT Modified-site 8 /note= "beta-Cyclohexylalanine"

WO200031129-A1.
 PN
 PD 02-JUN-2000.
 XX
 PP 24-NOV-1999; 99WO-EP009207.
 XX
 PR 26-NOV-1998; 98GB-00025946.
 XX

(RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX
 PI Pessi A, Ingallinella P, Bianchi E;
 XX
 DR WPI; 2000-411933/35.
 XX
 PT Peptide compounds useful for treating or preventing hepatitis C
 PT infections bind to hepatitis C virus NS3 protease.

XX Disclosure; Page 17; 46pp; English.
 XX
 CC The present sequence is one of the most preferred examples from a large
 CC number of peptides that were synthesised as potential inhibitors of
 CC hepatitis C virus NS3 serine protease. HCV infection can result in
 CC chronic hepatitis and cirrhosis of the liver, and may lead to
 CC hepatocellular carcinoma. Currently no vaccine nor established therapy
 CC exists. The present peptide is a reversible, non-covalent inhibitor that
 CC is based on the P and P' regions of the natural substrate of NS3. The P'
 CC part of the inhibitor is optimised to achieve maximum binding energy
 CC through interaction with the S' region of the enzyme. The peptides were
 CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
 CC The ability of the compounds to inhibit NS3 protease was evaluated using
 CC a complex comprising the NS3 protease domain and a modified form of the
 CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
 CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
 CC of the disclosed peptides to inhibit cleavage of the substrate by the
 CC complex was measured and some of the peptides were found to be potent
 CC inhibitors of HCV protease. The compounds can be used at micromolar and
 CC nanomolar levels to treat or prevent hepatitis C or a related condition
 CC
 CC Sequence 10 AA;

Query Match 92.5%; Score 37; DB 3; Length 10;
 Best Local Similarity 90.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
 | | | | |
 Db 1 DELIXCPXDL 10

RESULT 3
 AAY99544
 ID AAY99544 standard; peptide; 13 AA.
 AC AAY99544;
 XX
 DT 18-NOV-2004 (revised)
 DT 11-SEP-2000 (first entry)
 XX
 DE Hepatitis C virus NS3 protease peptidic inhibitor #88.
 XX
 KW Hepatitis C virus; HCV; serine protease; virucide; infection;
 KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
 KW hepatocellular carcinoma.
 XX
 OS Synthetic.

Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Misc-difference 2 /note= "D-form residue"
 FT Modified-site 5 /note= "beta-Cyclohexylalanine"
 FT Modified-site 8 /note= "beta-Cyclohexylalanine"
 FT Modified-site 13 /note= "N-epsilon-acetyl-lysine"

WO200031129-A1.
 PN
 PD 02-JUN-2000.
 XX
 PP 24-NOV-1999; 99WO-EP009207.
 XX
 PR 26-NOV-1998; 98GB-00025946.
 XX
 PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX
 PI Pessi A, Ingallinella P, Bianchi E;

XX WPI; 2000-411933/35.
DR Peptide compounds useful for treating or preventing hepatitis C
XX infections bind to hepatitis C virus NS3 protease.
PT Example 2; Page 32; 46pp; English.
XX
PS The present sequence is one of a large number of peptidic inhibitors of
XX hepatitis C virus NS3 serine protease. HCV infection can result in
CC chronic hepatitis and cirrhosis of the liver, and may lead to
CC hepatocellular carcinoma. Currently no vaccine nor established therapy
CC exists. The present peptide is a reversible, non-covalent inhibitor that
CC is based on the P and P' regions of the natural substrate of NS3. The P'
CC part of the inhibitor is optimised to achieve maximum binding energy
CC through interaction with the S' region of the enzyme. The peptides were
CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
CC The ability of the compounds to inhibit NS3 protease was evaluated using
CC a complex comprising the NS3 protease domain and a modified form of the
CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
CC of the disclosed peptides to inhibit cleavage of the substrate by the
CC complex was measured and some of the peptides were found to be potent
CC inhibitors of HCV protease. The present peptide is active at nanomolar
CC levels and could therefore be used to treat or prevent hepatitis C or a
CC related condition
CC
CC Revised record issued on 18-NOV-2004 : Missing qualifiers
XX Sequence 13 AA;
SQ
Query Match 92.5%; Score 37; DB 3; Length 13;
Best Local Similarity 90.0%; Pred. NO. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DYLIXCPXDL 10
DB 1 DELIXCPXDL 10
RESULT 4
AA99551
ID AA99551 standard; peptide; 13 AA.
AC AA99551;
XX
XX 18-NOV-2004 (revised)
DT 11-SEP-2000 (first entry)
DE Hepatitis C virus NS3 protease peptidic inhibitor #95.
XX
XX Hepatitis C virus; HCV; serine protease; virucide; infection;
KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
KW hepatocellular carcinoma.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT FT Misc-difference 2 /note= "D-form residue"
FT FT Misc-difference 5 /note= "beta-Cyclohexylalanine"
FT FT Modified-site 8 /note= "beta-Cyclohexylalanine"
FT FT Modified-site 13 /note= "N-Acetyl-lysine"
FT FT
XX WO200031129-A1.
XX
PN
XX
PD 02-JUN-2000.
XX

PF 24-NOV-1999; 99WO-EP009207.
XX
PR 26-NOV-1998; 98GB-00025946.
XX
PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
PI Pessi A, Ingallinella P, Bianchi E;
XX WPI; 2000-411933/35.
DR
XX Peptide compounds useful for treating or preventing hepatitis C
XX infections bind to hepatitis C virus NS3 protease.
XX Example 1; Page 24; 46pp; English.
XX
CC The present sequence is one of a large number of peptidic inhibitors of
CC hepatitis C virus NS3 serine protease. HCV infection can result in
CC chronic hepatitis and cirrhosis of the liver, and may lead to
CC hepatocellular carcinoma. Currently no vaccine nor established therapy
CC exists. The present peptide is a reversible, non-covalent inhibitor that
CC is based on the P and P' regions of the natural substrate of NS3. The P'
CC part of the inhibitor is optimised to achieve maximum binding energy
CC through interaction with the S' region of the enzyme. The peptides were
CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
CC The ability of the compounds to inhibit NS3 protease was evaluated using
CC a complex comprising the NS3 protease domain and a modified form of the
CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
CC of the disclosed peptides to inhibit cleavage of the substrate by the
CC complex was measured and some of the peptides were found to be potent
CC inhibitors of HCV protease
CC
CC Revised record issued on 18-NOV-2004 : Missing qualifiers
XX Sequence 13 AA;
SQ
Query Match 92.5%; Score 37; DB 3; Length 13;
Best Local Similarity 90.0%; Pred. NO. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DYLIXCPXDL 10
DB 1 DELIXCPXDL 10
RESULT 5
ADC42337
ID ADC42337 standard; protein; 1288 AA.
XX
XX ADC42337;
AC
XX 18-DEC-2003 (first entry)
DT
XX A. thaliana FANCD2 SEQ ID NO:3.
DE
XX
XX plant; cancer; Fanconi Anaemia; FA; BRCA; cytostatic; microarray;
KW chemosensitising.
KW
XX Arabidopsis thaliana.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 1.1288 /note= "X=Unknown"
FT
XX WO2003039327-A2.
PN
XX 15-MAY-2003.
PD
XX 06-JUN-2002; 2002WO-US018153.
PF
XX 02-NOV-2001; 2001US-00998027.
PR
XX 02-NOV-2001; 2001WO-US045561.
PR

PA (DAND) DANA FARBER CANCER INST.
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI D'andrea AD, Taniguchi T, Timmers C, Grompe M, Fox EA;
 XX
 DR WPI; 2003-441436/41.
 XX
 PT Diagnosing or determining cancer or increased risk of cancer in a
 PT patient, by testing Fanconi Anemia/BRCA pathway gene or protein for a
 PT cancer-associated defect, that indicates cancer or increased risk of
 PT cancer.
 XX
 PS Disclosure; SEQ ID NO 3; 160pp; English.
 XX
 CC The invention relates to a novel method of diagnosing or determining if a
 CC patient has cancer or is at increased risk of cancer, involving testing a
 CC Fanconi Anemia (FA)/BRCA pathway gene or protein for the presence of a
 CC cancer-associated defect, where the presence of one or more cancer-
 CC associated defects is indicative of cancer or an increased risk of cancer
 CC in the patient. The method of the invention has cytostatic activity. The
 CC method is useful for determining if a patient has cancer, or is at
 CC increased risk of developing cancer, e.g. breast, ovarian or prostate
 CC cancer. A microarray of the invention is useful for determining if a
 CC patient has cancer, or is at increased risk of developing cancer, by
 CC hybridising a nucleic acid sample to the nucleic acid sequences from the
 CC array, and detecting the presence of mutations in FA/BRCA pathway genes
 CC in the nucleic acid sample from the patient, where detecting the presence
 CC of mutations is indicative of a patient who has cancer, or is at
 CC increased risk of developing cancer. A method of the invention is useful
 CC for screening a chemosensitising agent, and the agent obtained is useful
 CC for treating a patient having a cancer. The present sequence is used in
 CC the exemplification of the invention.
 XX
 SQ Sequence 1288 AA;
 Query Match 90.0%; Score 36; DB 7; Length 1288;
 Best Local Similarity 60.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXLIXCPXDL 10
 | | | | |
 Db 228 DALLGCPLDL 237
 RESULT 6
 AAB93918
 ID AAB93918 standard; protein; 280 AA.
 XX
 AC AAB93918;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:13894.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 13894; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 280 AA;
 Query Match 85.0%; Score 34; DB 4; Length 280;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXLIXCPXD 9
 | | | | |
 Db 149 DSLIQCPID 157
 RESULT 7
 AAM38798
 ID AAM38798 standard; protein; 332 AA.
 XX
 AC AAM38798;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1943.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.

20-JUN-2000; 2000US-00598042.
19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI57954.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Example 3; SEQ ID NO 1943; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAI38642-AAI42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX Sequence 332 AA;
SQ
Query Match 85.0%; Score 34; DB 4; Length 332;
Beat Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXLIXCPXD 9
Db 286 DSLIQCPID 294
RESULT 8
AAU87274
ID AAU87274 standard; protein; 357 AA.
XX AAU87274;
AC AAU87274;
XX
DT 05-JUN-2002 (first entry)
XX
DE Novel central nervous system protein #184.
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
XX Homo sapiens.
XX
XX WO200155318-A2.
PN
XX

PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001332.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205551P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.

PN US2004018969-A1.
XX 29-JAN-2004.
PD 17-JAN-2001; 2001US-00764875.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 18-AUG-2000; 2000US-0225759P.
XX 22-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 26-SEP-2000; 2000US-0234998P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0255678P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2004-122079/12.
 DR N-PSDB; ADI53991.
 XX
 XX New polypeptides and nucleic acid molecules, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating or ameliorating medical
 PT conditions e.g. neural disorders, reproductive disorders or infectious
 PT diseases.
 XX
 XX Claim 11; SEQ ID NO 792; 413pp; English.
 PS
 CC The invention describes an isolated polypeptide comprising an amino acid
 CC sequence at least 90% identical to: a polypeptide fragment, domain,
 CC epitope, or full-length protein of any one of 607 amino acid sequences
 CC (I) described in the specification; a polypeptide fragment of (I), or the
 CC encoded sequence contained in (II), having biological activity; or a
 CC variant, allelic variant, or a species homologue of (I). The polypeptides
 CC and nucleic acid molecules are useful for detecting, preventing,
 CC diagnosing, prognosticating, treating or ameliorating medical conditions
 CC such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
 CC Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
 CC immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic

Query Match 85.0%; Score 34; DB 8; Length 357;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
 | | | | |
 Db 317 DSLIQCPID 325

RESULT 10
 ADI21120
 ID ADI21120 standard; protein; 399 AA.
 XX
 AC ADI21120;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Novel human protein #95.
 XX
 KW forensic; nutritional source; damaged tissue; diseased tissue;
 KW myeloid cell disorder; lymphoid cell disorder;
 KW bone cartilage tissue growth; tendon tissue growth;
 KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
 KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025148-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 19-SEP-2002; 2002WO-US029964.
 XX
 PR 19-SEP-2001; 2001US-0323739P.
 PR 13-SEP-2002; 2002US-00323739.
 XX
 XX (HYSE-) HYSEQ INC.
 FA
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
 PI Haley-Vicente D;
 XX
 XX WPI; 2003-354603/33.
 DR N-PSDB; ADI21836.
 DR
 XX New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and

PT replacement.
 XX
 PS Claim 20; SEQ ID NO 371; 156pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a
 CC polypeptide with biological activity. The polynucleotides and
 CC polypeptides are useful in diagnostics, forensics, gene mapping,
 CC identification of mutations responsible for genetic disorders and other
 CC traits, to assess biodiversity, as nutritional sources or supplements.
 CC The polynucleotides may also be used as molecular weight markers,
 CC chromosome markers or map related gene positions, or as an antigen to
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
 CC useful for raising antibodies, as markers for tissues in which the
 CC corresponding polypeptide is expressed, for re-engineering damaged or
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or
 CC regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. The
 CC present sequence represents the amino acid sequence of a novel human
 CC protein.
 XX
 SQ Sequence 399 AA;
 Query Match 85.0%; Score 34; DB 7; Length 399;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
 | | | | |
 Db 268 DSLIQCPID 276

RESULT 11
 AAY53029
 ID AAY53029 standard; protein; 417 AA.
 XX
 AC AAY53029;
 XX
 DT 29-FEB-2000 (first entry)
 XX
 DE Human secreted protein clone cw1640_1 protein sequence SEQ ID NO:64.
 XX
 KW Human; secreted protein; nutritional; cytokine; cell proliferation;
 KW differentiation; immune stimulating; vaccine; suppression;
 KW haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic;
 KW chemokinetic; haemostatic; thrombolytic; receptor; ligand;
 KW anti-inflammatory; cadherin; tumour invasion suppressor;
 KW tumour inhibition; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9957132-A1.
 XX
 PD 11-NOV-1999.
 XX
 PF 07-MAY-1999; 99WO-US009970.
 XX
 PR 07-MAY-1998; 98US-0084564P.
 PR 02-JUN-1998; 98US-0087645P.
 PR 22-JUL-1998; 98US-0093712P.
 PR 31-JUL-1998; 98US-0094935P.
 PR 10-AUG-1998; 98US-0095880P.
 PR 11-AUG-1998; 98US-0096068P.
 PR 06-MAY-1999; 99US-00306111.
 XX
 XX (GEMY) GENETICS INST INC.
 PA
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RU, Bowman MR;
 PI Diblasio-Smith E, Widom A;
 XX
 XX WPI; 2000-052937/04.
 DR N-PSDB; AAZ33347.
 DR

XX New polynucleotides encoding secreted human proteins, derived from adult
 PT placenta, adult retina, fetal brain, fetal.
 XX Claim 73; Page 418-419; 492pp; English.
 XX The present invention describes new human secreted proteins which were
 CC isolated from adult placenta, adult retina, foetal brain, foetal kidney,
 CC adult blood, adult brain, adult thyroid, adult bladder, adult neural
 CC tissue, adult testes, and adult lymph node cDNA libraries. The human
 CC secreted proteins, and the polynucleotides encoding them, are predicted
 CC to have biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals. Suggested activities include nutritional activity, cytokine and
 CC cell proliferation/differentiation activity, immune stimulating activity,
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The
 CC polynucleotides are also stated to be useful for gene therapy. AA233316
 CC to AA233373 encode human secreted proteins, and AA232998 to AA233060
 CC represent human secreted proteins, given in the present invention
 XX
 SQ Sequence 417 AA;

Query Match 85.0%; Score 34; DB 3; Length 417;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
 DB 286 DSIQCPID 294

RESULT 12
 AAB94044
 ID AAB94044 standard; protein; 417 AA.
 XX
 AC AAB94044;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:14203.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

PS Claim 8; SEQ ID NO 14203; 2537pp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification; where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB93893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 417 AA;

Query Match 85.0%; Score 34; DB 4; Length 417;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
 DB 286 DSIQCPID 294

RESULT 13
 AAM40584
 ID AAM40584 standard; protein; 437 AA.
 XX
 AC AAM40584;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 5515.
 XX
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-0052317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI59740.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 5515; 10078pp; English.
 CC
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with neurotropic, and the
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 437 AA;
 Query Match 85.0%; Score 34; DB 4; Length 437;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXD 9
 Db 306 DSLIQCPID 314
 RESULT 14
 ADI21597
 ID ADI21597 standard; protein; 437 AA.
 XX
 AC ADI21597;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Novel human polypeptide #76.
 XX
 KW forensic; nutritional source; damaged tissue; diseased tissue;
 KW myeloid cell disorder; lymphoid cell disorder;
 KW bone cartilage tissue growth; tendon tissue growth;
 KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
 KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
 XX
 OS Homo sapiens.
 XX
 FN WO2003025148-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 19-SEP-2002; 2002WO-US029964.
 XX
 PR 19-SEP-2001; 2001US-0323739P.
 PR 13-SEP-2002; 2002US-00323739.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
 PI Haley-Vicente D;
 XX WPI; 2003-354603/33.

DR N-PSDB; ADI21377.
 XX
 PT New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and
 PT replacement.
 XX
 PS Example 3; SEQ ID NO 848; 156pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a
 CC polypeptide with biological activity. The polynucleotides and
 CC polypeptides are useful in diagnostics, forensics, gene mapping,
 CC identification of mutations responsible for genetic disorders and other
 CC traits, to assess biodiversity, as nutritional sources or supplements.
 CC The polynucleotides may also be used as molecular weight markers,
 CC chromosome markers or map related gene positions, or as an antigen to
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
 CC useful for raising antibodies, as markers for tissues in which the
 CC corresponding polypeptide is expressed, for re-engineering damaged or
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or
 CC regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. The
 CC present sequence represents the amino acid sequence of a novel human
 CC polypeptide.
 XX
 SQ Sequence 437 AA;
 Query Match 85.0%; Score 34; DB 7; Length 437;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXD 9
 Db 306 DSLIQCPID 314
 RESULT 15
 AAU22976
 ID AAU22976 standard; protein; 440 AA.
 XX
 AC AAU22976;
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE Novel human enzyme polypeptide #62.
 XX
 KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.
 XX
 OS Homo sapiens.
 XX
 FN WO200155301-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001239.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.

CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 440 AA;

Query Match 85.0%; Score 34; DB 4; Length 440;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 DXLIXCPXD 9
Db 309 DSLIQCPID 317

Search completed: March 15, 2006, 19:44:25
Job time : 262.065 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:44:52 ; Search time 20.4348 Seconds
(without alignments)
47.085 Million cell updates/sec

Title: US-09-856-886B-104
Perfect score: 40
Sequence: 1 DXLIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Pirl1:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	82.5	294	2 T33588	hypothetical prote
2	33	82.5	816	2 H85028	probable GTP pyrop
3	33	82.5	1440	2 T27942	lin-15B protein -
4	32	80.0	98	2 JA0099	ferredoxin [2Fe-2S
5	32	80.0	218	2 G75420	hypothetical prote
6	32	80.0	487	2 T19875	hypothetical prote
7	32	80.0	677	2 T02951	probable mitogen a
8	32	80.0	697	2 C75525	molybdopterin oxid
9	31	77.5	553	2 D83640	hypothetical prote
10	30	75.0	278	2 S06173	development-specif
11	30	75.0	307	2 AF2216	inosine-adenosine-
12	30	75.0	320	2 G82293	lytB protein VC088
13	30	75.0	341	2 S57595	mitochondrial repl
14	30	75.0	418	2 T25092	hypothetical prote
15	30	75.0	446	2 E82835	ATP-dependent RNA
16	30	75.0	469	2 A64411	hypothetical prote
17	30	75.0	485	2 T04840	hypothetical prote
18	30	75.0	540	2 S70403	zona pellucida gly
19	30	75.0	683	1 Q0BEM8	gene 29 protein -
20	30	75.0	683	2 T42943	hypothetical prote
21	30	75.0	762	2 C69657	cobalamin-independ
22	30	75.0	830	2 T04848	protein kinase hom
23	29	72.5	53	2 F32946	cysteine proteinas
24	29	72.5	81	2 E81044	hypothetical prote
25	29	72.5	99	1 FEKT1	ferredoxin [2Fe-2S
26	29	72.5	126	1 FEKM	ferredoxin [2Fe-2S
27	29	72.5	182	2 D70859	hypothetical prote
28	29	72.5	224	2 H82395	transcription regu
29	29	72.5	257	2 T21029	hypothetical prote

30	29	72.5	270	2 T27786	hypothetical prote
31	29	72.5	300	2 AE0762	conserved hypothet
32	29	72.5	303	2 H70389	formate dehydrogen
33	29	72.5	303	2 H81395	probable ferrochel
34	29	72.5	314	2 T26438	hypothetical prote
35	29	72.5	337	2 A49970	glycine receptor a
36	29	72.5	339	2 A75138	meth2 homolog PAR0
37	29	72.5	351	2 AH1078	transcription regu
38	29	72.5	351	2 AG1436	transcription regu
39	29	72.5	373	2 C82701	lipoic acid synthe
40	29	72.5	377	2 S36081	probable carrier p
41	29	72.5	409	2 E70860	probable serB2 pro
42	29	72.5	527	2 S62484	hypothetical prote
43	29	72.5	548	2 S61096	probable membrane
44	29	72.5	639	2 D83591	probable ATP-depen
45	29	72.5	674	2 H72423	alpha-glucuronidas

ALIGNMENTS

RESULT 1

T33588
hypothetical protein C40A11.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33588
R:Maggi, L.; Goela, D.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid C40A11.
A:Reference number: Z21374
A:Accession: T33588
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <MAG>
A:Cross-references: UNIPROT:Q9T2B0; UNIPARC:UPI000007EEDC; EMBL:AF099914; PIDN:AAC68758.1
A:Experimental source: strain Bristol N2; clone C40A11
C:Genetics:
A:Gene: CESP:C40A11.4
A:Map position: 2 74/3, 103/3, 183/3, 222/1
A:Introns: 33/2, 74/3, 103/3, 183/3, 222/1

Query Match 82.5%; Score 33; DB 2; Length 294;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXLIXCPXD 9
DB 195 DELINCPED 203

RESULT 2

H85028
probable GTP pyrophosphokinase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H85028
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85028
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-816 <STO>
A:Cross-references: UNIPROT:Q9M112; UNIPARC:UPI00000488E2; GB:NC_001268; NID:G7268986; P1

Query Match 82.5%; Score 33; DB 2; Length 816;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
 Db 789 ESLVICPVDL 798

RESULT 3
 T27942
 lin-15B protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T27942; S47899; T27981
 R:White, S.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z20444
 A:Accession: T27942
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1440 <WIL>
 A:Cross-references: UNIPROT:Q27395; UNIPARC:UPI00000820CC; EMBL:Z79604; PIDN:CAB01901.1;
 A:Experimental source: clone ZK662
 R:Clark, S.G.; Lu, X.; Horvitz, H.R.
 Genetics 137, 987-997, 1994
 A:Title: The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine kin
 A:Reference number: S47893; MUID:95073612; PMID:7982579
 A:Accession: S47899
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1440 <CLA>
 A:Cross-references: UNIPARC:UPI00000820CC; EMBL:U10412; NID:G516015; PIDN:AAA20088.1; PI
 R:Kershaw, J.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z20449
 A:Accession: T27981
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1440 <WIL2>
 A:Cross-references: UNIPARC:UPI00000820CC; EMBL:Z79605; PIDN:CAB01908.1; GSPDB:GN000028;
 A:Experimental source: clone ZK678
 C:Genetics:
 A:Gene: CESP:ZK662.4
 A:Map position: X
 A:Introns: 25/2; 178/2; 185/3; 272/2; 304/3; 923/2; 1061/3; 1419/3

Query Match 82.5%; Score 33; DB 2; Length 1440;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
 Db 1013 DDILICPDL 1022

RESULT 4
 JA0099
 ferredoxin [2Fe-2S] - Ochromonas danica
 C:Species: Ochromonas danica
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C:Accession: JA0099
 R:Masui, R.; Wada, K.; Matsubara, H.; Williams, M.M.; Rogers, L.J.
 Phytochemistry 27, 2817-2820, 1988
 A:Title: Characterization, amino acid sequence and phylogenetic considerations regarding
 A:Reference number: JA0099
 A:Accession: JA0099
 A:Molecule type: protein
 A:Residues: 1-98 <MAS>
 A:Cross-references: UNIPROT:Q7M259; UNIPARC:UPI0000174DC9
 C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homolog
 C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein; metalloprotein
 F:26-80/Domain: ferredoxin [2Fe-2S] homolog <FER>
 F:41,46,49,79/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 80.0%; Score 32; DB 2; Length 98;

Best Local Similarity 55.6%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
 Db 15 DTIIDCPDD 23

RESULT 5
 G75420
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: G75420
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; P
 ; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75420
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <WHI>
 A:Cross-references: UNIPROT:Q9RV00; UNIPARC:UPI000000C18CF; GB:AE001371; GB:AE0000513; NID:
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRI230
 A:Map position: 1

Query Match 80.0%; Score 32; DB 2; Length 218;
 Best Local Similarity 55.6%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
 Db 37 DALLSCPPD 45

RESULT 6
 T19875
 hypothetical protein C41C4.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19875
 R:Burton, J.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: Z19190
 A:Accession: T19875
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-487 <WIL>
 A:Cross-references: UNIPARC:UPI0000179C86; EMBL:Z48045; PIDN:CAA88102.1; GSPDB:GN000020;
 A:Experimental source: clone C41C4
 C:Genetics:
 A:Gene: CESP:C41C4.7
 A:Map position: 2
 A:Introns: 45/2; 109/2; 189/3; 369/3; 446/3

Query Match 80.0%; Score 32; DB 2; Length 487;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
 Db 101 DGSITCPDL 110

RESULT 7
 T02951
 probable mitogen activated protein kinase - rice
 C:Species: Oryza sativa (rice)
 C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004

C;Accession: T02951
R;Kim, C.Y.; Cheon, S.Y.; Cho, M.J.
submitted to the EMBL Data Library, July 1998
A;Description: Identification and characterization of fungal elicitor responsive rice gene
A;Reference number: Z14788
A;Accession: T02951
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-877 <KIM>
A;Cross-references: UNIPROT:O81637; UNIPARC:UPI00000A443B; EMBL:AF080436; NID:G3450841;
A;Experimental source: cultivar Milliyang 117
C;Genetics:
A;Gene: MEK1

Query Match 80.0%; Score 32; DB 2; Length 677;
Best Local Similarity 55.6%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
| | | | |
Db 286 DDLVFCPGD 294

RESULT 8
C75525
molybdopterin oxidoreductase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004
C;Accession: C75525
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-697 <WHI>
A;Cross-references: UNIPROT:Q9RXB8; UNIPARC:UPI00000D3CB7; GB:AE001899; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0397
A;Map position: 1

Query Match 80.0%; Score 32; DB 2; Length 697;
Best Local Similarity 55.6%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
| | | | |
Db 29 DVLATCPDL 37

RESULT 9
D83640
hypoethetical protein PA0049 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: D83640
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83640
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-553 <STO>
A;Cross-references: UNIPROT:Q91783; UNIPARC:UPI00000C4F03; GB:AE004444; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:

A;Gene: PA0049

Query Match 77.5%; Score 31; DB 2; Length 553;
Best Local Similarity 44.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
| | | | |
Db 321 DGIVYCPAD 329

RESULT 10

S06173

development-specific protein D7 - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004

C;Accession: S06173

R;Smith, R.C.; Dworkin, M.B.; Dworkin-Rastl, E.

Genes Dev. 2, 1296-1306, 1988

A;Title: Destruction of a translationally controlled mRNA in Xenopus oocytes delays prog

A;Reference number: S06173; MUID:89079004; PMID:3203907

A;Accession: S06173

A;Molecule type: mRNA

A;Residues: 1-278 <SMI>

A;Cross-references: UNIPROT:P13007; UNIPARC:UPI0000128DC2; EMBL:X13856; NID:G64650; PIDN:

Query Match 75.0%; Score 30; DB 2; Length 278;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
| | | | |
Db 4 DELMQCPVD 12

RESULT 11

AF2216

inosine-adenosine-guanosine-nucleoside hydrolase [imported] - Nostoc sp. (strain PCC 7120

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AF2216

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AF2216

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-307 <KUB>

A;Cross-references: UNIPROT:Q8YS07; UNIPARC:UPI00000CE72C; GB:BA000019; PIDN:BA074984.1;

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr3285

C;Superfamily: yaaF protein

Query Match 75.0%; Score 30; DB 2; Length 307;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
| | | | |
Db 196 IIMCPLDL 203

RESULT 12

G82293

lytB protein VC0685 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: G82293

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Praser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82293
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <HEL>
A;Cross-references: UNIPARC:UPI0000164B5E; GB:AB004154; GB:AE003852; NID:g9655115; PIDN:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0685
A;Map position: 1
C;Superfamily: penicillin tolerance protein

Query Match 75.0%; Score 30; DB 2; Length 320;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
|||
Db 250 LIDCPQDI 257
|||

RESULT 13
S57595
mitochondrial replication protein MTF1 - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: protein RF1023; protein YM9959.10; protein YMR228w
C;Species: *Saccharomyces cerevisiae*
C;Date: 19-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S57595; S22842
R;Skellton, J.; Churcher, C.M.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57595
A;Accession: S57595
A;Molecule type: DNA
A;Residues: 1-341 <SKE>
A;Cross-references: UNIPROT:P14908; UNIPARC:UPI00001683E2; EMBL:Z49939; NID:g987599; PIDN:
A;Experimental source: Strain AB972
R;Liowski, T.; Michaelis, G.
Mol. Gen. Genet. 214, 218-223, 1988
A>Title: A nuclear gene essential for mitochondrial replication suppresses a defect of m
A;Reference number: S22842; MUID:89181520; PMID:2853292
A;Accession: S22842
A;Molecule type: DNA
A;Residues: 1-78,'S',80-286,'PI',289-341 <LIS>
A;Cross-references: UNIPARC:UPI0000168DDC; EMBL:X13513; NID:g4309; PIDN:CAA31864.1; PID:
C;Genetics:
A;Gene: SGD:MTF1
A;Cross-references: SGD:S0004841; MIPS:YMR228w
A;Map position: 13R
C;Superfamily: *Saccharomyces cerevisiae* mitochondrial replication protein MTF1
C;Keywords: mitochondrion

Query Match 75.0%; Score 30; DB 2; Length 341;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
|||
Db 297 DULKKCPIDL 306
|||

RESULT 14
T25092
hypothetical protein T22A3.3 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
R;McMurray, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19980

A;Accession: T25092
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-418 <WIL>
A;Cross-references: UNIPROT:P91820; UNIPARC:UPI000007BB36; EMBL:Z81125; PIDN:CAB03381.1;
A;Experimental source: clone T22A3
C;Genetics:
A;Gene: CESP:T22A3.3
A;Map position: 1
A;Introns: 24/3; 100/3; 123/3; 305/3; 390/2
C;Superfamily: *Caenorhabditis elegans* hypothetical protein T22A3.3

Query Match 75.0%; Score 30; DB 2; Length 418;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXD 9
|||
Db 16 LIACPSD 22
|||

RESULT 15
E82835
ATP-dependent RNA helicase XF0192 [imported] - *Xylella fastidiosa* (strain 9a5c)
C;Species: *Xylella fastidiosa*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82835
R;Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-446 <SIM>
A;Cross-references: UNIPROT:Q9PGV6; UNIPARC:UPI00000C231B; GB:AE003873; GB:AE003849; NID:
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briotes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.W.F.; Marino, C.L.; Marques, M.V.; Martins, E.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.F.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zs
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0192

Query Match 75.0%; Score 30; DB 2; Length 446;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
|||
Db 126 DLIACPGRL 135
|||

Search completed: March 15, 2006, 19:50:13
Job time : 21.4348 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:34:25 ; Search time 126.522 Seconds
(without alignments)
55.763 Million cell updates/sec

Title: US-09-856-886b-104
Perfect score: 40
Sequence: 1 DXLIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	90.0	403	2	Q4KA18_PSEF5
2	36	90.0	666	2	Q4QA30_LEIMA
3	36	90.0	2453	2	Q7RIN7_GIALA
4	34	85.0	219	2	Q86S77_HUMAN
5	34	85.0	417	1	ARP10_HUMAN
6	34	85.0	417	2	Q53HF6_HUMAN
7	34	85.0	417	2	Q53HF6_HUMAN
8	34	85.0	417	2	Q52ID3_CHICK
9	34	85.0	1299	2	Q7QXCL_GIALA
10	33	82.5	41	2	Q4X7V2_PLACH
11	33	82.5	265	2	Q9TZB0_CABEL
12	33	82.5	274	2	Q8ESZ6_OCBTH
13	33	82.5	351	2	Q61P98_CABER
14	33	82.5	404	2	Q5U563_XENLA
15	33	82.5	415	2	Q7ZVU0_BRARE
16	33	82.5	417	2	Q52KY3_XENLA
17	33	82.5	417	2	Q640T7_XENTR
18	33	82.5	417	2	Q6GNE3_XENLA
19	33	82.5	503	2	Q4WREL_ASPFU
20	33	82.5	583	2	Q585Q4_9TRYP
21	33	82.5	653	2	Q4QCT3_LEIMA
22	33	82.5	816	2	Q9M112_ARATH
23	33	82.5	1414	2	Q5XK85_XENLA
24	33	82.5	1440	2	Q27395_CABEL
25	32	80.0	71	2	Q6M9T2_PARUM
26	32	80.0	98	2	Q7M259_OCHDN
27	32	80.0	160	2	Q9N4A3_CABEL
28	32	80.0	179	2	Q61L22_DROME
29	32	80.0	210	2	Q5M6M4_CAMJE
30	32	80.0	218	2	Q9RY00_DEIRA
31	32	80.0	314	1	ISPH_VIBF1

32	32	80.0	333	2	Q9AD40_STRCO
33	32	80.0	374	2	Q8I4M4_CABEL
34	32	80.0	404	1	CTNS_CABEL
35	32	80.0	417	1	ARP10_MOUSE
36	32	80.0	417	2	Q5M9F7_RAT
37	32	80.0	448	2	Q6ZL25_ORISA
38	32	80.0	620	2	Q9SRR3_ARATH
39	32	80.0	677	2	Q81637_ORISA
40	32	80.0	697	2	Q9RXB8_DEIRA
41	32	80.0	703	2	Q6ZL26_ORISA
42	32	80.0	975	2	Q50R59_ENTHI
43	32	80.0	2539	2	Q9J7C1_PERSV
44	31	77.5	107	2	Q7MU06_PORGI
45	31	77.5	113	2	Q5LF95_BACFN

ALIGNMENTS

RESULT 1

Q4KA18_PSEF5
ID Q4KA18_PSEF5 PRELIMINARY; PRT; 403 AA.
AC Q4KA18_PSEF5 PRELIMINARY; PRT; 403 AA.
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Membrane protein, putative.
GN ORFNames=PFL3643;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-Pf-5;
RX PubMed=1598061; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.;
RT "complete genome sequence of the plant commensal Pseudomonas
fluorescens Pf-5.";
RL Nat. Biotechnol. 23:873-878 (2005).
DR EMBL; CP000076; AAY92909.1; -; Genomic DNA.
SQ SEQUENCE 403 AA; 44757 MW; A06BE49D4ABF0A2E CRC64;

Query Match 90.0%; Score 36; DB 2; Length 403;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXDL 10
DB 329 DALLICPQDL 338

RESULT 2

Q4QA30_LEIMA
ID Q4QA30_LEIMA PRELIMINARY; PRT; 666 AA.
AC Q4QA30;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=LmjP25.0650;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; C005264; CAJ04848.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 666 AA; 68968 MW; 7F092DBB6BC5B4CC CRC64;
 Query Match 90.0%; Score 36; DB 2; Length 666;
 Best Local Similarity 60.0%; Pred. No. 76;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXDL 10
 | | | | |
 Db 621 DVLLCPADL 630
 RESULT 3
 Q7R1N7 GIALA PRELIMINARY; PRT; 2453 AA.
 AC Q7R1N7
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GLP 28 44677 37316.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 OX NCBI_TaxID=184922;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAC801000026; EAA41222.1; -; Genomic DNA.
 SQ SEQUENCE 2453 AA; 275386 MW; F4032F7378971BEC CRC64;
 Query Match 90.0%; Score 36; DB 2; Length 2453;
 Best Local Similarity 60.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXDL 10
 | | | | |
 Db 1947 DFLHCPSDL 1956
 RESULT 4
 Q86SY7 HUMAN PRELIMINARY; PRT; 219 AA.
 AC Q86SY7
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Full-length cDNA clone CS0DJ008YD17 of T cells (Jurkat cell line) of
 DE Homo sapiens (human).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=T cells;
 RA Li W.B., Gruber C., Jessee J., Polayes D.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=T cells;
 RA Genoscope;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX161491; CAD61940.1; -; mRNA.

DR GO; GO:0015629; C:actin cytoskeleton; IEA.
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
 DR InterPro; IPR004000; Actin_like.
 DR Pfam; PF00022; Actin; 1.
 SQ SEQUENCE 219 AA; 24336 MW; 56354F3E00870959 CRC64;
 Query Match 85.0%; Score 34; DB 2; Length 219;
 Best Local Similarity 66.7%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 DXLIXCPXD 9
 | | | | |
 Db 88 DSLIQCPID 96
 RESULT 5
 ARP10 HUMAN STANDARD; PRT; 417 AA.
 AC Q9NZ32; O9H9Y5; O9NWX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Actin-related protein 10 (hARP11).
 GN Name=ACTR10; Synonyms=ACTR11, ARP11;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Hypothalamus;
 RA Xiao H., Song H., Gao G., Ren S., Chen Z., Han Z.;
 RL "A novel gene expressed in human hypothalamus.";
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Sano S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Ovary;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

Gene 138:171-174 (1994).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Brain;
 RC Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
 RA "Construction and characterization of a full length-enriched and a 5'-
 RT end-enriched cDNA library.";
 RL Gene 200:149-156 (1997).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Brain;
 RC Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
 RA Tanaka A., Yokoyama S.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the actin family.
 DR EMBL; AK222702; BAD96422.1; -; mRNA.
 DR GO; GO:0005515; F:protein binding; IEA.
 FT NON_TER 1
 SQ SEQUENCE 417 AA; 46307 MW; C57BB07F6737B0DA CRC64;

Query Match 85.0%; Score 34; DB 2; Length 417;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps

QY 1 DXLIXCPXD 9
 DB 286 DSLIQCPID 294

RESULT 7
 Q53HF6 HUMAN
 ID Q53HF6_HUMAN PRELIMINARY; PRT; 417 AA.
 AC Q53HF6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Uncharacterized hypothalamus protein HARP11 variant (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Cerebellum;
 RA Maruyama K., Sugano S.;
 RT "Oligo-capping : a simple method to replace the cap structure of
 RT eucaryotic mRNAs with oligoribonucleotides.";
 RL Gene 138:171-174 (1994).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Cerebellum;
 RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
 RT "Construction and characterization of a full length-enriched and a 5'-
 RT end-enriched cDNA library.";
 RL Gene 200:149-156 (1997).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Cerebellum;
 RC Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
 RA Tanaka A., Yokoyama S.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the actin family.
 DR EMBL; AK222624; BAD96344.1; -; mRNA.
 DR GO; GO:0005515; F:protein binding; IEA.
 FT NON_TER 1
 SQ SEQUENCE 417 AA; 46206 MW; E284AD0D91B10379 CRC64;

Query Match 85.0%; Score 34; DB 2; Length 417;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps

QY 1 DXLIXCPXD 9

```

Db          286 DSLIQCPCID 294
|||||
RESULT 8
Q52ID3_CHICK PRELIMINARY; PRT; 417 AA.
AC Q52ID3_
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORENAMES=RCJMB04.27n8;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=90311;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatki A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RA "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; AJ720851; CAG32510.1; -; mRNA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 2.
DR SMART; SM00268; ACTIN; 1.
DR KW Hypothetical protein.
SQ SEQUENCE 417 AA; 46281 MW; 0384C3A0716C4D22 CRC64;

Query Match 85.0%; Score 34; DB 2; Length 417;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy          1 DXLIXCPXD 9
|||||
Db          286 DSLIQCPCID 294

RESULT 9
Q7QXC1_GIALA PRELIMINARY; PRT; 1299 AA.
AC Q7QXC1_
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE GLP 741 38634 42533.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC801000061; EAA39702.1; -; Genomic_DNA.
SQ SEQUENCE 1299 AA; 143955 MW; 4020E9A5D3E0512A CRC64;

Query Match 85.0%; Score 34; DB 2; Length 1299;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy          1 DXLIXCPXD 10
|||||
Db          439 DMLMECPDLD 448

RESULT 10
Q4X7V2_PLACH PRELIMINARY; PRT; 41 AA.
AC Q4X7V2_
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORENAMES=PC405246.00.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN NUCLEOTIDE SEQUENCE.
RP Hall N., Karras M., Raine J.D., Carlton J.M., Koolij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RA "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAJ01009090; CAH87024.1; -; Genomic_DNA.
DR KW Hypothetical protein.
FT NON TER 41
SQ SEQUENCE 41 AA; 5147 MW; B51286E382674C11 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 41;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy          1 DXLIXCPXD 10
|||||
Db          13 DYLIRCPKEL 22

RESULT 11
Q9TZB0_CABEL PRELIMINARY; PRT; 265 AA.
AC Q9TZB0_
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein C40A11.4.
GN ORENAMES=C40A11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RT The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AF099914; AAC68758.2; -; Genomic_DNA.
DR PIR; T33588; T33588.
DR Ensembl; C40A11.4; Caenorhabditis elegans.
DR WormBase; WBGene00016547; C40A11.4.
DR WormPep; C40A11.4; CE36797.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008076; C:voltage-gated potassium channel complex; IEA.

```

DR GO: GO:0005515; F:protein binding; IEA.
 DR GO: GO:0005249; F:voltage-gated potassium channel activity; IEA.
 DR GO: GO:0006813; P:potassium ion transport; IEA.
 DR InterPro: IPR000210; BTP_POZ.
 DR InterPro: IPR003131; K_tetra.
 DR Pfam: PF02214; K_tetra; 1.
 DR PROSITE: PS00097; BTP; 1.
 KW Complete proteome; Hypothetical protein.
 DR PROSITE: PS00097; BTP; 1.
 SQ SEQUENCE 265 AA; 31019 MW; 758EF97197D3250D CRC64;

Query Match 82.5%; Score 33; DB 2; Length 265;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
 DB 166 DELINCPED 174

RESULT 12
 Q8ESZ6_OCEIH PRELIMINARY; PRT; 274 AA.
 AC Q8ESZ6;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical conserved protein.
 GN OrderedLocustNames=OB0469;
 OS Oceanobacillus theysensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; BA000028; BAC12425.1; -; Genomic DNA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0005058; P:biosynthesis; IEA.
 DR InterPro: IPR02060; Ssq/phyt_synthase.
 DR Pfam: PF00494; Ssq_Psr; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 274 AA; 31566 MW; 154A41BFD8BF7D42 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 274;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 10
 DB 102 DWLLFCPSDI 111

RESULT 13
 Q61P98_CABER PRELIMINARY; PRT; 351 AA.
 AC Q61P98;
 DT 25-OCT-2004 (TREMELrel. 28, Created)
 DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)
 DE Hypothetical protein CBG07639.
 GN Name=CBG07639;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

DR The C.briggsae Sequencing Consortium;
 DR GO: GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004000; Actin_like.

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CA001000032; CAB63265.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 351 AA; 39119 MW; 2A7E3FAF9B73EC55 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 351;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 10
 DB 23 DWIQCPCDL 32

RESULT 14
 Q5U563_XENLA PRELIMINARY; PRT; 404 AA.
 AC Q5U563;
 DT 01-FEB-2005 (TREMELrel. 29, Created)
 DT 01-FEB-2005 (TREMELrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
 DE LOC495357 protein (Fragment).
 GN Name=LOC495357;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Whole;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Whole;
 RA Klein S., Gerhard D.S.;
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the actin family.
 DR EMBL; BC084821; AAH84821.1; -; mRNA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR InterPro: IPR004000; Actin_like.

```
DR Pfam: PF00022; Actin: 1.
DR SMART: SMO0268; ACTIN: 1.
FT NON_TER 1
SQ SEQUENCE 404 AA; 44971 MW; EBB29A3AB197C4A7 CRC64;

Query Match      82.5%; Score 33; DB 2; Length 404;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 273 DSVQPCPD 281

RESULT 15
Q7ZVU0 BRARE
ID Q7ZVU0 BRARE PRELIMINARY; PRT; 415 AA.
AC Q7ZVU0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actin-related protein 10 homolog.
GN Name=actr10;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB; TISSUE=Whole body;
RG NIH MGC Project;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the actin family.
DR EMBL; BC045412; AAH45412.1; -; mRNA.
DR HSSP; P02568; 1KXP.
DR Ensembl; ENSDARG0000017663; Danio rerio.
DR ZFIN; ZDB-GENE-040426-768; actr10.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; Actin; 2.
DR SMART; SMO0268; ACTIN; 1.
SQ SEQUENCE 415 AA; 46006 MW; BD516BE11FDAE7DB CRC64;

Query Match      82.5%; Score 33; DB 2; Length 415;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
```

Db 284 DTLVKCPID 292

Search completed: March 15, 2006, 19:49:24
Job time : 129.522 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:49:41 ; Search time 28.6957 Seconds
(without alignments)
28.811 Million cell updates/sec

Title: US-09-856-886b-104
Perfect score: 40
Sequence: 1 DXLIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/aaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/aaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/aaa/ECTUS_COMB.pep.*
5: /cgn2_6/prodata/1/aaa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	85.0	159	2	US-09-621-976-4449
2	32	80.0	314	2	US-09-540-236-2699
3	32	80.0	677	2	US-09-345-473B-38
4	32	80.0	677	2	US-09-862-027-38
5	31	77.5	146	2	US-09-270-767-44634
6	31	77.5	347	2	US-09-540-236-3086
7	31	77.5	444	2	US-09-328-352-6463
8	31	77.5	596	2	US-09-543-681A-5487
9	21	77.5	601	2	US-09-252-991A-21824
10	30	75.0	15	2	US-09-894-998A-32
11	30	75.0	15	2	US-09-894-998A-33
12	30	75.0	15	2	US-10-237-551-32
13	30	75.0	15	2	US-10-237-551-31
14	30	75.0	19	2	US-09-441-502B-23
15	30	75.0	110	2	US-09-894-998A-22
16	30	75.0	110	2	US-10-237-551-22
17	30	75.0	147	2	US-09-894-998A-21
18	30	75.0	147	2	US-10-237-551-21
19	30	75.0	151	2	US-10-104-047-2708
20	30	75.0	318	2	US-09-894-998A-23
21	30	75.0	318	2	US-10-237-551-23
22	30	75.0	318	2	US-10-237-551-160
23	30	75.0	318	2	US-10-237-551-195
24	30	75.0	341	2	US-09-538-092-634
25	30	75.0	566	1	US-08-484-993B-41
26	30	75.0	566	1	US-08-484-158B-41
27	30	75.0	566	1	US-08-484-596A-41

28	30	75.0	566	1	US-08-480-150A-41	Sequence 41, Appl
29	30	75.0	566	2	US-08-458-731-41	Sequence 41, Appl
30	30	75.0	566	2	US-08-149-223A-41	Sequence 41, Appl
31	29	72.5	18	2	US-08-469-260A-385	Sequence 385, App
32	29	72.5	18	2	US-08-488-446-385	Sequence 385, App
33	29	72.5	18	2	US-08-467-344A-385	Sequence 385, App
34	29	72.5	18	2	US-08-424-550B-385	Sequence 385, App
35	29	72.5	68	2	US-09-270-767-38222	Sequence 38222, A
36	29	72.5	68	2	US-09-270-767-53439	Sequence 53439, A
37	29	72.5	79	1	US-08-839-709-3	Sequence 3, Appl
38	29	72.5	79	1	US-09-204-859-3	Sequence 3, Appl
39	29	72.5	126	2	US-09-900-920-55	Sequence 55, Appl
40	29	72.5	126	2	US-09-900-920-57	Sequence 57, Appl
41	29	72.5	246	2	US-09-134-001C-5070	Sequence 5070, Ap
42	29	72.5	307	2	US-09-489-039A-10116	Sequence 10116, A
43	29	72.5	375	2	US-09-270-767-45553	Sequence 45553, A
44	29	72.5	380	1	US-08-420-235B-5	Sequence 5, Appl
45	29	72.5	380	2	US-08-793-624-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-4449
; Sequence 4449, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4449
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4449

Query Match 85.0%; Score 34; DB 2; Length 159;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
DB 88 DSLIQCPID 96

RESULT 2

US-09-540-236-2699
; Sequence 2699, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2699
; LENGTH: 314
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2699

Query Match 80.0%; Score 32; DB 2; Length 314;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44634
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44634

Query Match          77.5%; Score 31; DB 2; Length 146;
Best Local Similarity 62.5%; Pred.No. 79;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DLIXCPXDL 10
   | : |||
Db 82 LLICPADL 89

RESULT 6
US-09-540-236-3086
; Sequence 3086, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3086
; LENGTH: 347
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-3086

Query Match          77.5%; Score 31; DB 2; Length 347;
Best Local Similarity 40.0%; Pred.No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 319 EALVGCPCVDI 328

RESULT 7
US-09-328-352-6463
; Sequence 6463, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6463
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6463

Query Match          77.5%; Score 31; DB 2; Length 444;
Best Local Similarity 50.0%; Pred.No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 416 EQLIECPIDI 425

RESULT 8
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44634
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44634

Query Match          77.5%; Score 31; DB 2; Length 146;
Best Local Similarity 62.5%; Pred.No. 79;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DLIXCPXDL 10
   | : |||
Db 82 LLICPADL 89

RESULT 6
US-09-540-236-3086
; Sequence 3086, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3086
; LENGTH: 347
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-3086

Query Match          77.5%; Score 31; DB 2; Length 347;
Best Local Similarity 40.0%; Pred.No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 319 EALVGCPCVDI 328

RESULT 7
US-09-328-352-6463
; Sequence 6463, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6463
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6463

Query Match          77.5%; Score 31; DB 2; Length 444;
Best Local Similarity 50.0%; Pred.No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 416 EQLIECPIDI 425

RESULT 8
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44634
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44634

Query Match          77.5%; Score 31; DB 2; Length 146;
Best Local Similarity 62.5%; Pred.No. 79;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DLIXCPXDL 10
   | : |||
Db 82 LLICPADL 89

RESULT 6
US-09-540-236-3086
; Sequence 3086, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3086
; LENGTH: 347
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-3086

Query Match          77.5%; Score 31; DB 2; Length 347;
Best Local Similarity 40.0%; Pred.No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 319 EALVGCPCVDI 328

RESULT 7
US-09-328-352-6463
; Sequence 6463, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6463
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6463

Query Match          77.5%; Score 31; DB 2; Length 444;
Best Local Similarity 50.0%; Pred.No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLIXCPXDL 10
   | : |||
Db 416 EQLIECPIDI 425

RESULT 8
```



```
US-09-543-681A-5487
; Sequence 5487, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5487
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5487

Query Match          77.5%; Score 31; DB 2; Length 596;
Best Local Similarity 60.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 DXLIXCPXDL 10
DB 454 DDLRDCPSDL 463

RESULT 9
US-09-252-991A-21824
; Sequence 21824, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21824
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (601)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-21824

Query Match          77.5%; Score 31; DB 2; Length 601;
Best Local Similarity 44.4%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
DB 369 DGIVYCPAD 377

RESULT 10
US-09-894-998A-32
; Sequence 32, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HSV-2
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-32

Query Match          75.0%; Score 30; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LIXCPXDL 10
DB 7 IVACPVDL 14

RESULT 11
US-09-894-998A-33
; Sequence 33, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-33

Query Match          75.0%; Score 30; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LIXCPXDL 10
DB 3 IVACPVDL 10

RESULT 12
US-10-237-551-32
; Sequence 32, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HSV-2
```

US-10-237-551-32

Query Match 75.0%; Score 30; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
: : |||
Db 7 IVACPVDL 14

RESULT 13

US-10-237-551-33

; Sequence 33, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HSV-2
US-10-237-551-33

Query Match 75.0%; Score 30; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
: : |||
Db 3 IVACPVDL 10

RESULT 14

US-09-441-502B-23

; Sequence 23, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-23

Query Match 75.0%; Score 30; DB 2; Length 19;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
: : |||
Db 9 LLKCPVDL 16

RESULT 15

US-09-894-998A-22

; Sequence 22, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:

; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-09-894-998A-22

Query Match 75.0%; Score 30; DB 2; Length 110;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
: : |||
Db 83 IVACPVDL 90

Search completed: March 15, 2006, 19:51:23

Job time : 29.6957 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2006, 19:50:26 ; Search time 98.6957 Seconds
(without alignments)
42.335 Million cell updates/sec

Title: US-09-856-886b-104
Perfect score: 40
Sequence: 1 DXLIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	85.0	51	4 US-10-425-115-216273	Sequence 216273,
2	34	85.0	357	3 US-09-764-875-792	Sequence 792, App
3	34	85.0	417	5 US-10-821-273-64	Sequence 64, Appl
4	34	85.0	440	3 US-09-764-853-594	Sequence 594, App
5	34	85.0	440	3 US-09-764-875-1102	Sequence 1102, App
6	34	85.0	442	3 US-09-764-875-809	Sequence 809, App
7	34	85.0	449	3 US-09-764-853-577	Sequence 577, App
8	33	82.5	32	4 US-10-437-963-164031	Sequence 164031,
9	33	82.5	914	5 US-10-450-763-44810	Sequence 44810, A
10	33	82.5	1102	5 US-10-450-763-38223	Sequence 38223, A
11	33	82.5	1102	5 US-10-450-763-44611	Sequence 44611, A
12	32	80.0	143	4 US-10-425-115-207095	Sequence 207095,
13	32	80.0	193	4 US-10-767-701-38761	Sequence 38761, A
14	32	80.0	333	4 US-10-282-122A-63074	Sequence 63074, A
15	32	80.0	677	3 US-09-862-027-38	Sequence 38, Appl
16	32	80.0	677	5 US-10-989-228-38	Sequence 38, Appl
17	32	80.0	678	4 US-10-437-963-183341	Sequence 183341,
18	32	80.0	697	4 US-10-369-493-23418	Sequence 23418, A
19	32	80.0	864	4 US-10-437-963-183335	Sequence 183335,
20	31	77.5	216	5 US-10-450-763-41280	Sequence 41280, A
21	31	77.5	298	4 US-10-389-566-565	Sequence 565, App
22	31	77.5	391	4 US-10-424-599-153298	Sequence 153298,
23	31	77.5	392	4 US-10-389-566-564	Sequence 564, App
24	31	77.5	396	4 US-10-389-566-525	Sequence 525, App
25	31	77.5	429	4 US-10-282-122A-63121	Sequence 63121, A
26	31	77.5	439	4 US-10-282-122A-45266	Sequence 45266, A
27	31	77.5	458	5 US-10-450-763-40267	Sequence 40267, A

28	31	77.5	946	4 US-10-437-963-111825	Sequence 111825,
29	30	75.0	15	3 US-09-894-998-32	Sequence 32, Appl
30	30	75.0	15	3 US-09-894-998-33	Sequence 33, Appl
31	30	75.0	15	4 US-10-121-988-32	Sequence 32, Appl
32	30	75.0	15	4 US-10-121-988-33	Sequence 33, Appl
33	30	75.0	15	4 US-10-200-562-32	Sequence 32, Appl
34	30	75.0	15	4 US-10-200-562-33	Sequence 33, Appl
35	30	75.0	15	4 US-10-237-551-32	Sequence 32, Appl
36	30	75.0	15	4 US-10-237-551-33	Sequence 33, Appl
37	30	75.0	15	5 US-10-945-050-32	Sequence 32, Appl
38	30	75.0	15	5 US-10-945-050-33	Sequence 33, Appl
39	30	75.0	92	4 US-10-424-599-276913	Sequence 276913,
40	30	75.0	110	3 US-09-894-998-22	Sequence 22, Appl
41	30	75.0	110	4 US-10-121-988-22	Sequence 22, Appl
42	30	75.0	110	4 US-10-200-562-22	Sequence 22, Appl
43	30	75.0	110	4 US-10-237-551-22	Sequence 22, Appl
44	30	75.0	110	5 US-10-945-050-22	Sequence 22, Appl
45	30	75.0	129	4 US-10-424-599-148473	Sequence 148473,

ALIGNMENTS

RESULT 1
US-10-425-115-216273
; Sequence 216273, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 216273
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_12883C.1.pep
US-10-425-115-216273

Query Match 85.0%; Score 34; DB 4; Length 51;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
DB 19 DSLIQCPID 27

RESULT 2
US-09-764-875-792
; Sequence 792, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 792
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-792

```
; NUMBER OF SEQ ID NOS: 939
; SEQ ID NO 594
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-594

Query Match      85.0%; Score 34; DB 3; Length 440;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 317 DSLIQCPID 325

RESULT 3
US-10-821-273-64
; Sequence 64, Application US/10821273
; Publication No. US20040248256A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: DiBiasio-Smith, Elizabeth
; APPLICANT: Widom, Angela
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000101.
; CURRENT APPLICATION NUMBER: US/10/821,273
; CURRENT FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: US 09/306,111
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: US 60/084,564
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: US 60/087,645
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: US 60/093,712
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: US 60/094,935
; PRIOR FILING DATE: 1998-07-31
; PRIOR APPLICATION NUMBER: US 60/095,880
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/096,068
; PRIOR FILING DATE: 1998-08-11
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-273-64

Query Match      85.0%; Score 34; DB 5; Length 417;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 286 DSLIQCPID 294

RESULT 4
US-09-764-853-594
; Sequence 594, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
```

```
; NUMBER OF SEQ ID NOS: 939
; SEQ ID NO 594
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-594

Query Match      85.0%; Score 34; DB 3; Length 440;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 309 DSLIQCPID 317

RESULT 5
US-09-764-875-1102
; Sequence 1102, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-1102

Query Match      85.0%; Score 34; DB 3; Length 440;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 309 DSLIQCPID 317

RESULT 6
US-09-764-875-809
; Sequence 809, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 809
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-809

Query Match      85.0%; Score 34; DB 3; Length 442;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXD 9
Db 311 DSLIQCPID 319
```

RESULT 7

US-09-764-853-577
 ; Sequence 577, Application US/09764853
 ; Patent No. US20020090672A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PJ206
 ; CURRENT APPLICATION NUMBER: US/09/764,853
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 939
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 577
 ; LENGTH: 449
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (15)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-764-853-577

Query Match 85.0%; Score 34; DB 3; Length 449;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
 |||||
 Db 318 DSLIQCPID 326

RESULT 8

US-10-437-963-164031
 ; Sequence 164031, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 164031
 ; LENGTH: 32
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_62971C.1.pep
 ; US-10-437-963-164031

Query Match 82.5%; Score 33; DB 4; Length 32;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 10
 |||||
 Db 20 DFLIGCPNLS 29

RESULT 9

US-10-450-763-44810
 ; Sequence 44810, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 44810
 ; LENGTH: 914
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (543)..(589)
 ; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
 ; OTHER INFORMATION: eMATRIX, accession number DM013540, p-value=1.000e-40, raw score
 ; OTHER INFORMATION: 8.73
 ; US-10-450-763-44810

Query Match 82.5%; Score 33; DB 5; Length 914;
 Best Local Similarity 50.0%; Pred. No. 5.9e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 10
 |||||
 Db 66 DHLVLCPDV 75

RESULT 10

US-10-450-763-38223
 ; Sequence 38223, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 38223
 ; LENGTH: 1102
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (543)..(589)
 ; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
 ; OTHER INFORMATION: eMATRIX, accession number DM013540, p-values=1.000e-40, raw score
 ; OTHER INFORMATION: 8.73
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (576)..(801)
 ; OTHER INFORMATION: Reverse transcriptase domain identified by Pfam, accession
 ; OTHER INFORMATION: name rvt, E-value=5.9e-48, Pfam score of 172.8
 ; US-10-450-763-38223

Query Match 82.5%; Score 33; DB 5; Length 1102;
 Best Local Similarity 50.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
| | | | |
Db 66 DHLVLCPODV 75

RESULT 11

US-10-450-763-44611
; Sequence 44611, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CI3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44611
; LENGTH: 1102
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: {543}..(589)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM013540, p-value=1.000e-40, raw score
; OTHER INFORMATION: 8.73
US-10-450-763-44611

Query Match 82.5%; Score 33; DB 5; Length 1102;
Best Local Similarity 50.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
| | | | |
Db 66 DHLVLCPODV 75

RESULT 12

US-10-425-115-207095
; Sequence 207095, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 207095
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_12045C.1.pgp
US-10-425-115-207095

Query Match 80.0%; Score 32; DB 4; Length 143;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LIXCPXDL 10
| | | | |

Db 113 LIKCPMDL 120

RESULT 13

US-10-767-701-38761
; Sequence 38761, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 38761
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(193)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C79162_1.pgp
US-10-767-701-38761

Query Match 80.0%; Score 32; DB 4; Length 193;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
| | | | |
Db 184 DCLLCPVDM 193

RESULT 14

US-10-282-122A-63074
; Sequence 63074, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

Thu Mar 16 08:05:40 2006

; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 63074
 ; LENGTH: 333
 ; TYPE: PRT
 ; ORGANISM: Moraxella catarrhalis
 US-10-282-122A-63074

Query Match 80.0%; Score 32; DB 4; Length 333;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 10
 DB 247 DTIYQPSDL 256

RESULT 15
 US-09-862-027-38
 ; Sequence 38, Application US/09862027
 ; Patent No. US2002014248A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hodge, Martin R.
 ; TITLE OF INVENTION: No. US2002014248A1 Kinases and Uses Thereof
 ; FILE REFERENCE: 35800/234862
 ; CURRENT APPLICATION NUMBER: US/09/862,027
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/345,473
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 38
 ; LENGTH: 677
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-862-027-38

Query Match 80.0%; Score 32; DB 3; Length 677;
 Best Local Similarity 55.6%; Pred. No. 7e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
 DB 286 DDLVFCFGD 294

Search completed: March 15, 2006, 19:55:16
 Job time : 99.6956 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run On: March 15, 2006, 19:51:36 ; Search time 13.0435 Seconds
(without alignments)
21.944 Million cell updates/sec

Title: US-09-856-886B-104
Perfect score: 40
Sequence: 1 DLIIXCPXDL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/prodata/1/pubpaa/US08 NEW PUB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US06 NEW PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US07 NEW PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US09 NEW PUB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10 NEW PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11 NEW PUB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US12 NEW PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US13 NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	82.5	158	7	US-11-098-686-11309
2	31	77.5	311	7	US-11-096-568A-3250
3	31	77.5	311	7	US-11-096-568A-3251
4	31	77.5	317	7	US-11-096-568A-3249
5	31	77.5	318	7	US-11-096-568A-3248
6	30	75.0	151	7	US-11-072-512-2708
7	30	75.0	156	7	US-11-096-568A-25076
8	30	75.0	181	7	US-11-120-308-68
9	30	75.0	208	7	US-11-096-568A-15370
10	30	75.0	306	7	US-11-096-568A-15369
11	30	75.0	320	7	US-11-096-568A-15368
12	30	75.0	532	7	US-11-120-308-72
13	29	72.5	98	6	US-10-510-386-138
14	29	72.5	144	7	US-11-124-367A-266
15	29	72.5	339	7	US-11-124-367A-265
16	29	72.5	352	7	US-11-124-367A-267
17	29	72.5	1189	6	US-10-821-234-1209
18	28	70.0	57	7	US-11-242-553-14
19	28	70.0	110	7	US-11-072-512-3717
20	28	70.0	143	7	US-11-096-568A-19073
21	28	70.0	164	7	US-11-087-099-1249
22	28	70.0	219	7	US-11-096-568A-19072
23	28	70.0	222	7	US-11-096-568A-19071
24	28	70.0	301	7	US-11-072-175-248
25	28	70.0	326	7	US-11-096-568A-22130

26	28	70.0	332	7	US-11-096-568A-26863	Sequence 26863, A
27	28	70.0	373	7	US-11-096-568A-22129	Sequence 22129, A
28	28	70.0	377	7	US-11-096-568A-26862	Sequence 26862, A
29	28	70.0	410	7	US-11-096-568A-12704	Sequence 12704, A
30	28	70.0	412	7	US-11-096-568A-12704	Sequence 22128, A
31	28	70.0	426	6	US-10-467-657-4808	Sequence 4808, Ap
32	28	70.0	537	7	US-11-096-568A-12703	Sequence 12703, A
33	28	70.0	591	7	US-11-096-568A-12702	Sequence 12702, A
34	28	70.0	759	7	US-11-096-568A-29706	Sequence 29706, A
35	28	70.0	764	7	US-11-096-568A-29705	Sequence 29705, A
36	28	70.0	767	7	US-11-096-568A-29704	Sequence 29704, A
37	28	70.0	972	7	US-11-177-894-17	Sequence 17, Appl
38	28	70.0	976	7	US-11-148-770-31	Sequence 31, Appl
39	28	70.0	976	7	US-11-177-894-15	Sequence 15, Appl
40	28	70.0	976	7	US-11-177-894-16	Sequence 16, Appl
41	28	70.0	976	7	US-11-177-894-18	Sequence 18, Appl
42	28	70.0	976	7	US-11-177-894-19	Sequence 19, Appl
43	28	70.0	976	7	US-11-177-894-20	Sequence 20, Appl
44	28	70.0	976	7	US-11-177-894-21	Sequence 21, Appl
45	28	70.0	988	7	US-11-207-078-112	Sequence 112, Appl

ALIGNMENTS

RESULT 1
US-11-098-686-11309
; Sequence 11309, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11309
; LENGTH: 158
; TYPE: PRT
; ORGANISM: lawsonia intracellularis
US-11-098-686-11309

Query Match 82.5%; Score 33; DB 7; Length 158;
Best Local Similarity 75.0%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LIXCPXDL 10
|||
Db 96 LIACPHDL 103

RESULT 2
US-11-096-568A-3250
; Sequence 3250, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3250
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Glycine max

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(311)
OTHER INFORMATION: Ceres Seq. ID no. 15171166
US-11-096-568A-3250

Query Match 77.5%; Score 31; DB 7; Length 311;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
|::|||:
Db 153 DCMLLCPADM 162

RESULT 3

US-11-096-568A-3251
Sequence 3251, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 3251
LENGTH: 311
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(311)
OTHER INFORMATION: Ceres Seq. ID no. 16625542
US-11-096-568A-3251

Query Match 77.5%; Score 31; DB 7; Length 311;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
|::|||:
Db 153 DCMLLCPADM 162

RESULT 4

US-11-096-568A-3249
Sequence 3249, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 3249
LENGTH: 317
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(317)
OTHER INFORMATION: Ceres Seq. ID no. 15171165
US-11-096-568A-3249

Query Match 77.5%; Score 31; DB 7; Length 317;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10

Db 159 DCMLLCPADM 168
|::|||:

RESULT 5

US-11-096-568A-3248
Sequence 3248, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 3248
LENGTH: 318
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(318)
OTHER INFORMATION: Ceres Seq. ID no. 15171164
US-11-096-568A-3248

Query Match 77.5%; Score 31; DB 7; Length 318;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DXLIXCPXDL 10
|::|||:
Db 160 DCMLLCPADM 169

RESULT 6

US-11-072-512-2708
Sequence 2708, Application US/11072512
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2708
LENGTH: 151
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-2708

```

Query Match          75.0%; Score 30; DB 7; Length 151;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 DXLIXCPXDL 10
DB 61 DKLSQCPWDL 70

RESULT 7
US-11-096-568A-25076
; Sequence 25076, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 25076
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(156)
; OTHER INFORMATION: Ceres Seq. ID no. 12556308
US-11-096-568A-25076

Query Match          75.0%; Score 30; DB 7; Length 156;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 4; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 1 DXLIXCPXDL 10
DB 37 DCMLLCPVDM 46

RESULT 8
US-11-120-308-68
; Sequence 68, Application US/11120308
; Publication No. US20060005277A1
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Forge, Charlie
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: cDNAs Encoding Polypeptides
; FILE REFERENCE: BB-1365 US NA
; CURRENT APPLICATION NUMBER: US/11/120,308
; CURRENT FILING DATE: 2005-05-02
; PRIOR APPLICATION NUMBER: US/10/078,770
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/614,188
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,400
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/153,534
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 60/161,223
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/159,878
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/157,401
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/143,419
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,409
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Microsoft Office 97

```

```

; SEQ ID NO 68
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Zea mays
US-11-120-308-68

Query Match          75.0%; Score 30; DB 7; Length 181;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 3 LIXCPXD 9
DB 86 LISCTPD 92

RESULT 9
US-11-096-568A-15370
; Sequence 15370, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 15370
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(208)
; OTHER INFORMATION: Ceres Seq. ID no. 12344353
US-11-096-568A-15370

Query Match          75.0%; Score 30; DB 7; Length 208;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 4; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 1 DXLIXCPXDL 10
DB 51 DCMLLCPVDM 60

RESULT 10
US-11-096-568A-15369
; Sequence 15369, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 15369
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(306)
; OTHER INFORMATION: Ceres Seq. ID no. 12344352
US-11-096-568A-15369

Query Match          75.0%; Score 30; DB 7; Length 306;
Best Local Similarity 40.0%; Pred. No. 41;
Matches 4; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 1 DXLIXCPXDL 10

```


RESULT 15
US-11-124-367A-265
; Sequence 265, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 265
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-367A-265

Query Match 72.5%; Score 29; DB 7; Length 339;
Best Local Similarity 44.4%; Pred. NO. 73;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXLIXCPXD 9
DB 53 DEWICPYD 61

Search completed: March 15, 2006, 19:55:51
Job time : 13.0435 secs

This Page Blank (uspto)